Sorry to Burst Your Bubble!

Developing an Educational Application to Raise Awareness About Filter Bubbles Among Teenagers

Tim van de Wiel
5972450

Master thesis
Human Computer Interaction
Utrecht University

19 / 07 / 2021

First examiner
Dr. Paweł W. Woźniak

Second examiner
Prof. dr. ir. Judith F. M. Masthoff

ABSTRACT

The algorithmic personalisation of online information can cause users to end up in so-called filter bubbles, which limit exposure to opposing viewpoints. Dutch teenagers are not aware of this online personalisation, and schools struggle to educate them on this. To help teachers and educators address the effects and dangers of online personalisation, the anti-filter bubble application project was established by an interdisciplinary research team. In this study, employing an iterative user-centred approach by interviewing students and teachers, I developed and tested two activities and an overall structure for the application. Additionally, I performed an evaluation study using one of the prototypes, investigating the influence of the presence of a physical artefact on knowledge gains and students' experience of that activity. Results showed that while students and teachers have competing interests, the presence of a physical object is likely to improve the subsequent class discussion. Along with the prototypes, recommendations for future development, and the evaluation study, I also contribute recommendations for future HCI research and developments with teenagers.

CONTENTS

1. INTRODUCTION 5
2. RELATED WORK 6
2.1 Project background 6
2.2 Information personalisation and the filter bubble 7
2.3 Teenagers 10
2.4 Human-Computer Interaction for Teenagers 13
3. PRE-STUDY 1: REQUIREMENTS 18
3.1 Interviews 19
3.2 Context of use 20
3.3 Requirements 22
3.4 Storyboarding 26
4. PRE-STUDY 2: STORYBOARDING 29
4.1 Methodology 29
4.2 Results 30
4.3 Discussion 32
5. DESIGN 34
5.1 Methods 34
5.2 Activity 1 39
5.3 Activity 2 43
5.4 Activity 3 55
5.5 Overall application structure 64
5.6 Summary 70
6. EVALUATION 71
6.1 Method for students 72
6.2 Method for teachers 77
6.3 Evaluation results 78
6.4 Evaluation discussion 86
7. DISCUSSION 90
7.1 Recommendations for future development 90
7.2 Limitations 95
7.3 Conclusion 96
ACKNOWLEDGEMENTS 98
REFERENCES 99

Appendix A. Interview protocol for pre-study 1 105
Introduction 105
Questions 105
Additional list of topics 105
Appendix B. Insights from pre-study 1 107
Appendix C. Interview protocol and topic list pre-study 2 109
Appendix D. Insights from pre-study 2 111
Appendix E. Design phase: students participants table 113
Appendix F. Design phase: teacher interview protocol 114
Appendix G. Design phase: students protocol 116
Appendix H. Activity 2 videos 119
H.1 First iteration 119
H.2 Second iteration 119
H.3 Third iteration 119
H.4 Fourth iteration 120
Appendix I. Activity 3 storylines 121
I.1 Jayden 121
I.2 Sophie 122
Appendix J. Evaluation questionnaires 124
J.1 Game Experience Questionnaire (GEQ) 124
J.2 Dispositional Flow Scale (DFS-2) 124
J.3 Filter bubble knowledge questions 125
Appendix K. Evaluation protocol and topic list for teachers 126

1. INTRODUCTION

Dutch secondary schools are sounding the alarm about the increasing segregation and polarisation in their classes, which they are largely attributing to the online environment their students find themselves in. In that online environment, especially on social media, algorithmic personalisation of information causes teenagers to be exposed only to content that aligns with their interests and attitudes. Although seemingly innocent and convenient, information personalisation restricts students' access to other viewpoints which further reinforces their opinions and in turn prevents the understanding of others. This is a direct cause for concern as it leads to a growing intolerance for others and propagates the development of more extreme and polarized attitudes (Pariser, 2011).

Teenagers in these schools are generally unaware of the personalisation algorithms at play online and the filter bubbles they are in, and schools and teachers are struggling with finding a way to educate teenagers on this particular topic. Previously developed tools to aid the teacher in this regard are unfit because they are either not developed for teenagers specifically, not tested nor evaluated in the real world, outdated, not suitable for the Dutch context or just not effective (Amrollahi, 2019). Given the continuous increase of polarisation among teenagers, there is an urgency in providing educators with a suitable tool to support them in their lessons on this topic.

Therefore, at the initiative of Utrecht University and in cooperation with Mira Media[1] and a select number of schools in Utrecht, the interdisciplinary anti-filter bubble project was established. The goal of this project is to develop a web-application to be used in class that raises teenagers' awareness of filter bubbles and provides them with the knowledge and skills needed to fight these filter bubbles. In this thesis, I will employ an iterative user-centred design approach to design prototypes of the application and evaluate them with the teenagers that will eventually be using the application. Through this design process I will also contribute both methodological and design recommendations and requirements to the field of Human-Computer Interaction with teenagers, which has received little attention thus far. By answering the following research question, future HCI studies involving teenagers or on awareness will be able to build on the recommendations brought forward by this thesis:

RQ *What are the design qualities of interactive tools that effectively raise awareness among teenagers?*

The next chapter contains related work on filter bubbles, teenagers and HCI with teenagers. Then, two pre-studies are conducted which present the context of use, requirements and the first storyboards. The fifth chapter is on the design process, and details all the steps and iterations that the prototypes have gone through. Following the design chapter is the evaluation study on one of the prototypes, and finally, Chapter 7 contains the discussion and future recommendations.

[1] Now transformed into *Het Mediateam*

2. RELATED WORK

This chapter provides an overview of the relevant areas for this project and is divided into four parts. The first section details the background of the anti-filter bubble project, how it came to be and what this thesis contributes to the project in general. Second, to be able to understand why bursting filter bubbles is important, information personalisation is discussed along with its dangers and proposed methods to counteract it. Third, understanding the user is vital for user-centred design. Section 2.3 is therefore about the characteristics and needs of teenagers and their relation to technology and polarisation. Finally, the current work on HCI for teenagers is reviewed in order to gain an understanding of how previous studies have approached designing with and for teenagers, as that allows the current study to build on the methods and techniques used in those studies.

2.1 Project background

Together with the municipality of Utrecht and several secondary schools in Utrecht, Utrecht University is part of the UNION[2] network which aims to prevent polarisation in those secondary schools. The anti-filter bubble project was conceived after secondary school teachers in Utrecht indicated that they were concerned about the role of social media in their classes and the contribution it has to polarisation. The project is embedded in Utrecht University's *Change Your Perspective* interdisciplinary research hub, allowing researchers from different research fields to contribute their expertise to the project. Scholars and students from Pedagogy, Computer Science, Information Sciences, Human-Computer Interaction and Sociology are all involved in the project.

The goal of the anti-filter bubble project is twofold: first, it will result in an educational application to be used in class. The application teaches teenagers about their filter bubble, provides them with methods to subvert bubbles and it can be integrated into secondary school lesson plans. To this end, five secondary schools in Utrecht are directly involved in the project. Second, the project will contribute to the scientific knowledge about (awareness of) filter bubbles and the effectiveness of anti-bubble software in education.

The project kicked off in 2019. With the UNION network and 'digital social broker' Mira Media as acting clients, bachelor students in Information Sciences developed early rudimentary prototypes for an educational application relying on the basic principles of user-centred design. Based on these prototypes, a group of bachelor students in Computer Sciences and Game Art worked together in 2020 in a software project that resulted in a further advanced prototype. Two master students in Youth, Education & Society (YES) assisted in developing pedagogical interventions to be included in the application and delivered master theses on teenagers' appreciation of filter bubbles.

This thesis is part of a new phase of the anti-filter bubble project. In this phase, the (continued) development of the anti-filter bubble application is the main focus. Together with two new YES-students (Zena Bani and Anouk Adriani) and one other HCI-student (Anneleen Janssen),

[2] Utrechts Netwerk voor Inclusie in het ONderwijs; *Utrecht Network for Inclusion in Education*

secondary school students and teachers are involved closer to the development of the application. Employing an iterative design cycle will allow for continuous refinement of the prototype and eventually result in a final prototype that is adequately evaluated.

2.2 Information personalisation and the filter bubble

The last two decades have seen a dramatic increase in the amount of online information available for everyone (Geschke, Lorenz, & Holtz, 2018). This means it is not possible to gather, read and interpret all accessible content, and a selection has to be made. Zuiderveen Borgesius et al. (2016) describe how this information personalisation can be accomplished in two different ways. First, *self-selected personalisation* involves users selecting information themselves which is usually information closely aligned to their own interests and opinions (Burbach, Halbach, Ziefle, & Valdez, 2019; Zuiderveen Borgesius et al., 2016). Information that contradicts or challenges their beliefs is usually avoided. This phenomenon is also referred to as selective exposure (Burbach et al., 2019; Spohr, 2017).

The second type of information personalisation is *pre-selected personalisation* (Zuiderveen Borgesius et al., 2016). This type of personalisation is directed by recommender systems that filter online information based on the user's interests, demographics, and other features (Burbach et al., 2019; Amrollahi, 2019). Aside from recommender systems in online stores recommending similar products to the ones previously bought, pre-selected personalisation is most common on social media (Sindermann, Elhai, Moshagen, & Montag, 2020). On YouTube for example, recently watched videos dictate its recommendations (O'Callaghan, Greene, Conway, Carthy, & Cunningham, 2013) and search results (Hussein, Juneja, & Mitra, 2020). Facebook arranges the feed in such a way that it displays the posts most likely to be appreciated by a user on top of the page (Rader & Gray, 2015). While this type of personalisation might sound appealing to users, there are adverse consequences to it discussed below.

2.2.1 Filter bubbles and polarisation

A hypothesized outcome of pre-selected personalisation is the emergence of a filter bubble (Pariser, 2011). In such a "unique universe of information" (Pariser, 2011, p. 10), users are only presented with information or opinions that match with their own views (Burbach et al., 2019). Pariser describes three characteristics of filter bubbles. First, they are individual: no two users share the same experience. Second, a filter bubble is invisible as users are not notified about the recommendation systems. Even if users are familiar with the concept of a filter bubble, they are often either not fully aware of the extent of personalisation (Burbach et al., 2019) or assume that the underlying algorithms are unbiased (Rader & Gray, 2015). Finally, entering or creating the filter bubble is not a conscious choice of the user. Taken together, Pariser argues that a filter bubble "fundamentally alters the way we encounter ideas and information" (Pariser, 2011, p. 10). For example, YouTube recommends politically charged videos to their users that are related to their previously watched videos, pushing them further into more extreme ideological directions (Bahara, Kranenberg, & Tokmetzis, 2019; Bryant, 2020). Facebook filters out posts from friends or news outlets that do not align with your political preferences (Chitra & Musco, 2020; Pariser, 2011).

This fundamental alteration thus limits exposure to opposing viewpoints on political or moral issues unbeknownst to the users (Bozdag & Van der Hoven, 2015; Nechushtai & Lewis, 2019). The resulting gap in knowledge, or the "incomplete picture of the world" as Nechushtai & Lewis (2019, p. 300) remark, prevents a mutual understanding and restricts sensemaking between different ideological groups (Bozdag & Van der Hoven, 2015). On an individual level, this growing intolerance for other perspectives combined with the continuous reinforcement of similar beliefs cause the development of more extreme and more polarized attitudes (Courtois, Slechten, & Coenen, 2018; Stroud, 2008; Stroud, 2010). This leads to widespread polarisation and encourages radicalisation on a societal level (Amrollahi, 2019; Geschke et al., 2018), with multiple scholars warning about the adverse consequences for democracy (Bozdag & Van der Hoven, 2015; Flaxman, Goel, & Rao, 2016; Stroud, 2008; Zuiderveen Borgesius et al., 2016).

Some authors have argued that the premise of a filter bubble caused by algorithmic personalisation is incorrect, or at least not as detrimental to society as mentioned above. They suggest that the individual actions and attitudes of users are the root cause of the selective exposure they experience on social media (Bruns, 2019; Mahrt, 2019; Seargeant & Tagg, 2018). Individual attributes certainly do have an influence: Sindermann et al. (2020) for example showed that age, gender, and certain aspects of personality all correlate with the number of news sources one consumes. Additionally, a personal interest in politics or current affairs and a feeling of duty to keep informed are considerable factors in exposure to other viewpoints (Mahrt, 2019). The general trend in contemporary research however is that personalisation algorithms do at least partially cause the formation of filter bubbles.

Even as debates on the direct causes and origins of filter bubbles continue, its role in limiting individuals' exposure to opposing viewpoints is undeniable. Its contribution to polarisation in society cannot be understated, and the need for a strategy to combat filter bubbles is clear.

2.2.2 Bursting the filter bubble

There is a lack of research on users' awareness of filter bubbles (Gran, Booth, & Bucher, 2020; Plettenberg et al., 2020), and the small number of studies that are available are not conclusive in their findings. Nevertheless, recent studies do hint towards a generally low awareness of personalisation algorithms and filter bubbles (Gran, Booth, & Bucher, 2020; Hitlin & Rainie, 2019; Powers, 2017). Plettenberg et al. (2020) and Powers (2017) also noted that awareness was highest in those using social media the most. However, high awareness of algorithmic personalisation and filter bubbles does not seem to motivate users to take avoiding action (Plettenberg et al., 2020; Rader & Gray, 2015). Pariser (2011) originally suggested preventing filter bubbles through 'sabotaging' personalisation algorithms by deleting browsing history, using incognito mode or rejecting cookies. Unfortunately, these methods are tedious, ineffective and rarely performed by users (Bozdag & Van der Hoven, 2015; Burbach et al., 2019). Most of the users however do "express the wish for a tool helping them to do so [break their filter bubble]" (Plettenberg et al., 2020, p. 89).

Amrollahi (2019) identifies two different categories of such possible tools. The first category includes those that help users identify the filter bubble they are in, and how this bubble affects them. Milan & Agosti (2019) argue that tools like these foster *algorithm literacy* and are key to

improving filter bubble awareness. They can for example show users the posts that they missed out on because of the personalisation algorithm, or show them the news feeds from different users (see for example TheirTube[3]). The second category involves tools that aim to break the filter bubble. Within this category, two approaches can be distinguished (Resnick et al., 2013). One relatively straightforward approach is to provide users with more diverse information than they would have gotten in their bubble. News aggregators for example could show the highest rated counter-attitudinal articles (Resnick et al., 2013). The second approach is to subtly encourage and motivate users to search articles that do not match with their attitudes and preferences themselves. Providing users with feedback about their reading behaviour is already enough to nudge them towards a more balanced consumption of news (Resnick et al., 2013).

Multiple implementations of tools have already been proposed. For example, Nagulendra & Vassileva (2014) proposed an interactive visualisation that allows a user to directly control the filtering algorithms. Munson, Lee & Resnick (2013) developed a browser widget, *Balancer*, showing users the political (im)balance of their reading behaviour. Fouquaert (2019) produced *Instawareness,* a tool allowing Instagram users to see the impact of personalisation algorithms on the composition of their feed. Chrysanthou, Barlas, Kyriakou, Kleanthous & Otterbacher (2020) prototyped a demonstration system for educational purposes, illustrating the effects of personalisation in search engines. However, an overall review of most tools by Amrollahi (2019) shows a "lack of empirical studies on the effectiveness of the proposed tools" (p. 17). Tools are rarely tested, and if they are, testing is most often done on hypothetical platforms built specifically for the evaluation, limiting the generalizability of results to actual online platforms. Based on his review and conclusions, Amrollahi also proposes that future anti-filter bubble tools should not just focus on one but comprise both categories mentioned above: identify and alert the user about the filter bubble they are in, and help them burst that bubble.

Bozdag & Van der Hoven (2015) also note that most tools are developed within the political context of the United States, which is systematically different to most other Western countries. More empirical experiments in real online platforms need to be conducted in different political contexts and cultures (Amrollahi, 2019; Bozdag & Van der Hoven, 2015), and "breaking bubbles requires an interdisciplinary approach" (Bozdag & Van der Hoven, 2015, p. 263).

2.2.3 Summary

Filter bubbles emerge through online information personalisation, be it either self-selected or pre-selected, often on social media. Users of these social media should be aware of the dangers of filter bubbles as they limit exposure to other perspectives without notice, prevent understanding of other points of view and potentially increase polarized attitudes. Unfortunately, users are rarely fully aware of filter bubbles. Even if they are, possible strategies to break their bubble are still unknown to them. Those users do however wish for a tool that can help them fight their bubble. Such a tool could help unaware users identify the bubble they are in or provide them with opportunities to break it. Despite the potential there is a lack of tools that are actually developed, evaluated and implemented, let alone within a Dutch context. This highlights the importance of the current study, in which a prototype for such a tool is developed

[3] https://www.their.tube

and evaluated in close collaboration with the actual end users. In this tool, users need to be both be made aware about the filter bubble they are in and be taught the knowledge they need to break their filter bubble. Earlier studies have primarily focused on either one of those but rarely combined the two approaches.

2.3 Teenagers

The anti-filter bubble application is aimed at teenagers aged 12 to 14 enrolled in vmbo (*voorbereidend middelbaar beroepsonderwijs,* pre-vocational education*).* There are more than 200.000 students in vmbo, making it the biggest educational type in the Netherlands (Centraal Bureau voor de Statistiek [CBS], 2020a). This section provides a short introduction on teenagers, their relationship with online technology and approaches for preventing polarisation among them.

Grasping the characteristics, attitudes and needs of the entire population of teenagers has proven to be nearly impossible. They are not only a very diverse and multifaceted group, but are also rapidly developing both cognitively and physically (Bell, 2016; Rose, Björling, Kim, & Alvarez, 2018). Compared to younger children for example they are better able to reflect on their thinking (Rose et al., 2018), have greater control over their attention (Bell, 2016), show an increase in their working and long-term memory skills and can process information considerably quicker (Bell, 2016; Fitton & Bell, 2014). However, it must also be noted that the cognitive abilities of teenagers are not yet equal to those of adults, especially when it comes to decision making (Bell, 2016). Teenagers are more sensitive to social influences, take more risks and have an increased desire for self-expression and individuality (Bell, Fitton, & Toth, 2013; Fitton et al., 2016).

Joyce & Nielsen (2019) have identified three aspects in which teenagers especially differ from adults: (1) their reading skills are lower, (2) they use less sophisticated research strategies and (3) they have less patience. These three aspects need to be considered when working with or designing for teenagers - for example by relying less on textual instructions. As Rose et al. (2018) emphasize, however, "teens are extremely diverse, individualized, and highly contextualized, making it very challenging to generalize their opinions and preferences" (p. 2). User studies on the specific target group for this study are therefore a necessity.

2.3.1 Teenagers and technology

As the current study aims to introduce a technological tool that influences teenagers' online behaviour or at least raises awareness about the risks of filter bubbles, it is important to examine the relationship teenagers have with technology. Technology has become omnipresent in the daily lives of teenagers. In 2019, 99% of teenagers in the Netherlands had access to a personal mobile phone and 94% had a laptop at their disposal (CBS, 2020b). Additionally, a growing number of schools have their students use laptops, tablets, or online learning platforms such as Google Classroom (DiFranzo et al., 2019; Huisman, 2020). As teenagers today were already born in a technologically saturated world and are online for more than four hours a day on average (Rose et al., 2018; Overbeek & Van Rijn-van Gelderen, 2019), one could assume they can easily deal with the challenges and risks that online environments pose to them. With over 90% of

Dutch teenagers possessing more than the basic skills required to navigate modern technology (CBS, 2020c), it seems appropriate to call them *digital natives* (Kirschner & De Bruyckere, 2017).

The opposite is true, however; empirical evidence shows that teenagers' technological skills are often overestimated (Bower, 2017; Kirschner & De Bruyckere, 2017). They struggle with cyberbullying, cannot ensure their own online privacy, and perhaps most importantly, they have trouble distinguishing false information from the truth (DiFranzo et al., 2019; Dennen, Choy, & Word, 2020; Nieuwelink, 2020). Especially as searching for information and reading news articles online are considered low-risk activities by teenagers (Byrne et al., 2016), it is important to improve their awareness of the algorithmic personalisation that could shape these activities (Radicalisation Awareness Network [RAN], 2018).

This algorithmic personalisation is most prevalent on social media, which is proving to be a breeding ground for filter bubbles (Sindermann et al., 2020). Simultaneously, using social media is the most popular activity on the internet for teenagers by some distance. Nearly 95% of Dutch teenagers on the internet make use of social media (CBS, 2020b). Even though previous studies have been conducted on teenagers' behaviour on social media, the digital landscape they find themselves in is changing fast. Previous studies have often used Facebook as a focal point for their research (e.g. Rader & Gray, 2015; Seargeant & Tagg, 2018; Spohr, 2017), but teenagers are no longer using Facebook en masse. Instead, it has become only the fourth most popular social medium among Dutch teenagers aged 12-14 (Newcom, 2020). YouTube is most frequently used, followed by Instagram and Snapchat. Meanwhile, the relatively new medium TikTok is on the rise in terms of user base size with over 155 thousand 12– to 14-year-olds using the app in 2019. It is important to keep in mind that these usage statistics are constantly changing, and the tool developed in this project should be designed to accommodate this dynamic nature of social media. While using contemporary examples of social media will be more relatable for the current generation of teenagers, it might also cause the tool to become obsolete in the near future.

2.3.2 Teenagers and polarisation

Increasing polarisation among teenagers is not just a hypothetical scenario. Platform JEP[4], a Dutch governmental institution working to combat extremism and polarisation among teenagers, says that groups of teenagers have increasingly polarized attitudes against one another (JEP, 2019). They are not only feeling disconnected from other teenagers, but from Dutch society as a whole. The contrasts are especially visible between groups that are different in ethnicity or religion, and teenagers with a non-Western migration background are feeling increasingly excluded and distanced from the rest of society (JEP, 2019). As nearly 27% of vmbo students have a non-Western migration background (CBS, 2020a), this is particularly relevant for the current study.

Teenagers are more susceptible to polarising influences than others (JEP, 2019). They are still developing their sense of identity and those who struggle with that seek comfort within groups of like-minded people (JEP, 2019; Nederlands Jeugdinstituut[5] [NJI], 2018; Mutsaers & Demir,

[4] *Platform Jeugd preventie Extremisme en Polarisatie*; Platform for Youth prevention Extremism and Polarisation
[5] Netherlands Youth Institute

2020). Compared to adults like parents or teachers, peers have the most influence on teenagers (Bell, 2016; Shin & Lwin, 2017; RAN, 2018) which strengthens groupthink and provokes an 'us vs. them' mentality (JEP, 2019). Algorithmic personalisation on social media and the subsequent filter bubbles allows this groupthink and 'us vs. them' mentality to develop even further.

Preventing polarisation among teenagers is complicated. Youth workers, teachers and parents rarely possess the knowledge or skills available to deal with the increasingly polarized attitudes of teenagers (JEP, 2019; Mutsaers & Demir, 2020; Van Bergen, De Ruyter, & Pels, 2017). Platform JEP (2019) provides a number of starting points for prevention approaches. Aside from helping teenagers to find their place in society through jobs and communal work, which is outside the scope of this project, they also stress the need for teaching them the skills and knowledge needed to function in a multicultural society. Nieuwelink (2020) articulated this knowledge into three components: realizing that (1) societal issues will always involve highly different perspectives; (2) tolerating other perspectives is important; and (3) sharing opinions and attitudes with others is necessary. Perhaps most importantly however, both Platform JEP (2019) and the European Radicalisation Awareness Network (RAN, 2018) also emphasize the importance of involving teenagers in the process of developing interventions itself. Providing teenagers with an active role in development is already a step towards more resilience to polarisation.

Multiple scholars also underline the importance of schools and other educational institutes taking up a serious role in preventing polarisation (DiFranzo et al., 2019; Kirschner & De Bruyckere, 2017; RAN, 2018; Shin & Lwin, 2017). If children grow up in their own culture separated from others, high schools will be the first place where they meet other 'bubbles' (JEP, 2019). Teenagers are most influenced by their peers and schools are the best place to foster discussions between peers. Debates and discussions on sensitive topics between peers at school can help remove barriers among students and help students accept other perspectives (RAN, 2018). The educational tool developed in this project can capitalize on that by facilitating and fostering discussions between teenagers.

Additionally, RAN (2018) recommends improving *digital literacy* (or *media literacy*) among teenagers as a means for preventing polarisation. Being digitally literate means not just being able to handle technological devices, but also means being able to critically examine online information (Thijs, Fisser, & Van der Hoeven, 2014) and recognize any political or ideological bias in that information (Curriculum.nu, 2019). Schools are increasingly acknowledging the value of teaching digital literacy and it is becoming a part of curricula of schools worldwide, including that of the Netherlands in the near future (Curriculum.nu, 2019; Thijs, Fisser, & Van der Hoeven, 2014). Teaching digital literacy at schools helps teenagers develop "crucial skills and competencies for the use of (social) media" (RAN, 2018, p. 14) and "involves raising awareness and understanding [...] of the technical functions and algorithms that contribute to the visibility of related content" (RAN, 2018, p.14), thus providing them with the skills and knowledge necessary to be able to deal with the consequences of information personalisation. As the aim of the anti-filter bubble application developed in this thesis is to raise teenagers' awareness of filter bubbles and what they can do about them, it implicitly improves digital literacy.

2.3.3 Summary

Teenagers are a unique population, cognitively different to younger children and adults, with very diverse characteristics and needs. They have a strong relationship with technology, as nearly all of them own a personal mobile phone and are on average online for more than four hours a day. That does not mean they can be presumed to be *digital natives* however, and their technological and online skills are often overestimated. In any potential tool developed for teenagers, individual differences in knowledge and skills on online behaviour should thus be accounted for.

Individual differences are also reflected in the increasing polarisation among teenagers in the Netherlands. Groups of teenagers are feeling increasingly removed from society or other groups of teenagers, especially when these groups differ in ethnicity or religion. This is especially the case for teenagers with a non-Western migration background (making up more than a quarter of all vmbo students) who feel excluded from the rest of society. Preventing or decreasing polarisation among teenagers is complicated, as parents, teachers and youth workers are often unsure about what to do. Multiple scholars and experts advocate for a bigger role for schools and educational institutions in preventing polarisation by fostering discussions between teenagers and improving their digital literacy. The tool developed in this thesis can provide support for this by facilitating these discussions and teaching teenagers the skills and competencies needed for using digital media.

2.4 Human-Computer Interaction for Teenagers

Until surprisingly recently, teenagers in human-computer interaction (HCI) were treated either like young adults or old children (Fitton et al., 2016). Yarosh, Radu, Hunter, & Rosenbaum (2011) were among the first to draw attention to the lack of HCI research focusing specifically on teenagers. Seven years later, Rose et al. (2018) noted that "teen-computer interaction" was still an underrepresented area within HCI. This slow development might have had logistic or pragmatic reasons such as accessibility difficulties (Read & Horton, 2016). Nevertheless, teenagers should not be overlooked: as described in the previous section, they are avid users of technology and HCI should also cater to their needs. This section describes some of the considerations and challenges that come with HCI for teenagers.

2.4.1 Considerations

When designing for and with teenagers, some aspects that are unique to that demographic should be considered. There is not one agreed upon set of guidelines, but I will discuss four major topics that are essential in HCI research with teenagers based on several previous studies.

2.4.1.1 Group dynamics

Working with a group of teenagers in for example focus groups can make them more comfortable with sharing their thoughts and opinions than in an individual setting (Fitton et al., 2016). However, the way the group is composed (i.e. who are in the group) has a sizable influence on the discussions that are held. Large age differences between group members might cause younger teenagers to feel intimidated by the older teenagers, and for those older teenagers

mixed-gender groupings can cause some to be more constrained in their responses (Fitton et al., 2016; Raby, 2010). When a group is composed of teenagers that are friends with one another, they challenge other's ideas more often, speak more freely, are more willing to engage in a discussion and are more creative (Fitton et al., 2016; Horkoff, Ersare, Kahler, Jörundsson, & Hammouda, 2018; Raby, 2010). Group size is also of importance: a group that is too big prevents shyer teenagers from engaging with the discussions. Based on previous research and their experiences, Fitton et al. (2016) recommend a group size of 4-7 to spark balanced discussions that allow everyone to get a word in.

Wilkinson (2016) emphasizes that the usual sampling guidelines in HCI should also be adhered to when composing groups of teenagers. This means that groups should preferably consist of teenagers with a broad range of backgrounds and abilities. Minorities and teenagers from marginalized groups should also be included (Fitton et al., 2016), especially given the topic of polarisation in the current study.

2.4.1.2 Informality

To get and keep teenagers engaged with the research activities, two notions are especially important. First, teenagers are not children and therefore should not be treated as such. Childish design activities or immature use of language are certain to disengage teenagers from the task at hand (Fitton et al., 2016). This does not however imply that teenagers are to be treated as adults. Design activities should be both playful and serious (Fitton et al., 2016), but individual differences can produce a preference to either.

Second, there is inevitably a skewed power relationship between the researchers and the participating teenagers. Especially when within a school setting, teenagers will relate the researchers with their teachers as authoritative figures (Fitton, Read, & Horton, 2013; Poole & Peyton, 2013). This power relationship has to be balanced to get the most out of any activity. To accomplish this, design activities should be kept casual and light-hearted (Horkoff et al., 2018; Iivari, Kinnula, Kuure, & Keisanen, 2020), easy to understand and entertaining (Rose et al., 2018). Perhaps most importantly however, Iversen & Smith (2012) stress that "evaluators and other stakeholders must put aside their assumptions of superiority based on age and cognitive maturity" (p. 113).

2.4.1.3 Motivation and rewards

Rewards for participating need to be considered to maximize engagement (Fitton, Read, & Horton, 2013). Extrinsic rewards such as sweets or vouchers are useful in recruiting teenagers, creating enthusiasm among participants and they help to signal that the design activities are not a school situation (Fitton et al., 2016; Iversen, Dindler, & Hansen, 2013). However, extrinsic rewards only are not sustainable for long-term research projects and should not be the sole motivation for participating (Hansen & Iversen, 2013; Iversen, Dindler, & Hansen, 2013). It should also be rewarding to just take part: the activities themselves have to be engaging enough (Fitton et al., 2016). Endorsing the teenage participants as stakeholders who genuinely can have an impact or influence on design choices greatly increases motivation and engagement for longer periods of time (Hansen & Iversen, 2013; Iversen, Dindler, & Hansen, 2013; Kinnula & Iivari, 2019).

2.4.1.4 Techniques and methods

A review by Fitton et al. (2016) demonstrated that most traditional HCI research methods, like interviews and focus groups, "appear to be successful in engaging the teen population and producing high quality HCI research and design outputs" (p. 264). They do need to be adjusted appropriately. For example, broad open-ended questions in interviews can be difficult to answer for teenagers (Poole & Peyton, 2013). Focus groups can make teenagers more comfortable, but special attention needs to be given to the group composition. Questionnaires need to be easily understandable, for example by making use of emoji to simplify Likert scales (Bell, 2007; Horkoff et al., 2018). Other common research methods such as storyboarding or observational studies have been proven useful as well when working with teenagers (Horkoff et al., 2018), and some have adapted methods that are used in child-computer interaction to be suitable for teenagers (Fitton et al., 2016).

If possible, using multiple techniques throughout the research project or even within one session helps sustain the engagement of teenagers (Fitton et al., 2016; Raby, 2010). Finally, as Fitton et al. (2016) underline, there is not a 'one size fits all' technique in any study (p. 241). Specific contexts require specific research techniques. Individual differences between teenagers also impact the preferred method as some teenagers would rather complete questionnaires or written tasks than engage in playful activities (Fitton et al., 2016).

2.4.2 Educational tools

Aside from understanding how to do HCI research with teenagers, previous HCI studies have also examined or developed tools for education and can provide key recommendations to explore when designing the current application. The anti-filter bubble application is far from being the first technological tool intended for use in education. In the past few years, technology has changed education, and it will continue to change the way teenagers are learning in the future (Escutea, Nickow, Oreopoulos, & Quan, 2020). Online tools, educational games and digital interventions are currently commonplace in classrooms. They have long been implemented in traditional classes like mathematics (e.g. Roschelle et al., 2010) and science (e.g. Roschelle, Penuel, Yarnall, Schechtman, & Tatar, 2005), but also in (digital) citizenship education (e.g. Maqsood, Mekhail, & Chiasson, 2018; Yap & Lee, 2020). Overall, *technology for learning* has become a research field of its own in recent times and cannot be covered completely within the scope of this thesis. The following paragraphs will therefore examine the types of educational tools for related topics currently available and important considerations for developing such tools.

DiFranzo et al. (2019) have identified three categories of currently available educational tools to improve digital literacy. The first of these categories is *non-interactive curricula*, which do not allow for individual interaction between students and the topic at hand. They consist of predetermined scripted lesson plans or activities. The second kind of tools are *unguided platforms* or *sandboxes*, which are not necessarily educational tools by itself, but rather specifically designed platforms for a young audience. These platforms can give young students hands-on experiences with online material in a safe environment. The third category is the most relevant for the current study: *interactive edutainment 'games'*. These are often narrative-based tools where the student takes on the role of a character navigating hypothetical scenarios.

The interactive edutainment tools are "designed on the premise that high engagement through interactive gaming assists in learning and retention" (DiFranzo et al., 2019, p. 2). They can range from simple minigames with limited connection to real-world situations, to fully fleshed out serious games (DiFranzo et al., 2019). Regardless of their granularity, edutainment tools in general have been proven to be effective in supporting learning and motivating students because it makes learning more 'familiar' to them (Bower, 2017; Escutea et al., 2020). Unfortunately, most interactive tools in education developed thus far are fun for one-time use but do not help develop any meaningful skills in the long-term (DiFranzo et al., 2019). It is important to realize that after the novelty effect subsides, technology in and of itself will not suddenly engage students with the content. Excitement about technology is not the same as engagement with the learning goals (Kolb, 2017). The design of the learning task itself is paramount and technology should merely be seen as a means to enable additional ways of implementing learning tasks (Bower, 2017). Kolb (2017) advises to not only consider student engagement, but also to acknowledge *enhancement* (what does technology add that traditional tools cannot) and *extension* (how well technology transfers classroom knowledge to everyday life). Emphasizing engagement only might cause the users to not be able to draw the relations between in-game activities and their real-world implications (Maqsood, Mekhail, & Chiasson, 2018; Yap & Lee, 2020).

There are other factors that also influence the effectiveness and usefulness of educational technology. I have for example previously mentioned the need for individuality among teenagers and the importance of adjusting design methods to individual differences. In the same notion, Jin & Divitini (2020) underline the impact of individual differences in affinity towards technology. Finally, the crucial role of a teacher cannot be understated. A review by Bower (2017) found that "[the] role of the teacher was perhaps the most pervasive theme throughout all of the technology-enhanced learning literature" (p. 418), and Shin & Lwin (2017) mention that teachers can have a substantial impact on the online behaviour of teenagers. Teachers should be guiding the learning process, and their actions and instructions ultimately determine whether lessons with technology are a success (Bower, 2017; Kolb, 2017). Involving teachers should therefore be a priority in designing educational technologies, as that will help to implement any system effectively in their classrooms (Tatar, Roschelle, Vahey, & Penuel, 2003).

2.4.3 Summary

In HCI, teenagers are often treated as old children or young adults, while they should be an entirely unique population. Designing with and for them requires special considerations to be made. Special attention should be given to the composition of a group of teenagers in focus groups, informality during design sessions, and teenagers' motivation for participating. Design methodologies should be adapted to the specific group of teenagers that is studied.

Previous work on developing educational tools for teenagers, some with the same goal of improving digital literacy, has provided three categories of educational tools. Non-interactive curricula do not allow for interactivity, even though having a personal influence is important for teenagers. Unguided platforms do provide personalized hands-on experiences, but are less suitable for in-class usage and are not necessarily informative. The tool developed in this thesis would therefore likely be in between those categories: an interactive edutainment game. They have proven to be effective in learning, but the design of the learning content itself remains

paramount. Additionally, multiple studies have underlined the importance of the role of the teacher in guiding the learning process. In the tool developed in this thesis, this role should be well defined and adequately integrated into the tool itself by directly involving teachers in the design process.

3. PRE-STUDY 1: REQUIREMENTS

In order to gain insight into the context of use of the anti-filter bubble application, we performed an empirical preliminary study on the requirements of the application. This chapter describes that study and establishes the first requirements. The first section describes the methodology for this study, after which Section 2 delineates the context of use and Section 3 provides the requirements based on that context of use. Finally, Section 4 introduces the storyboards that were drafted based on the findings of this first preliminary study.

Figure 1 below displays the setup of the two pre-studies.

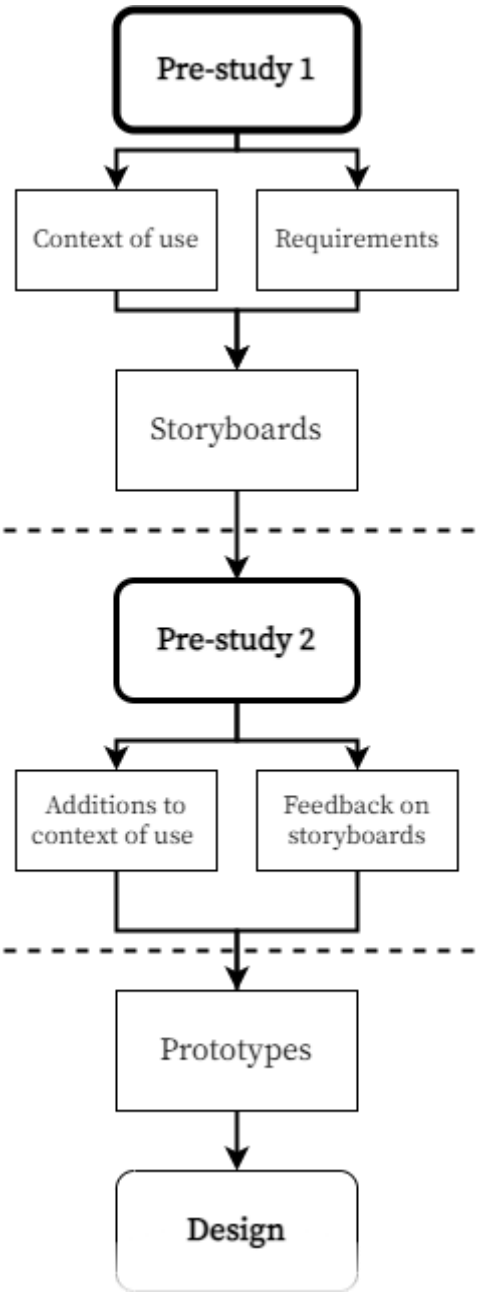

Figure 1: The setup of the two pre-studies. Rectangles correspond to the (sub)sections of the pre-study's chapters.

3.1 Interviews

To gain an understanding of the context in which the application will be used, we carried out semi-structured interviews with teachers in secondary schools. The goal of these interviews was to let the teachers tell us about their experiences with technology in class, how their students behave online, and how they would approach teaching and discussing topics like polarisation and filter bubbles in their classes. We developed questions beforehand (see Appendix A for the full protocol), but the format of a semi-structured interview provides some margin for omitting planned questions or for additional follow-up questions when necessary.

3.1.1 Participants

Five teachers and one expert in the field of educational technologies took part in this first study. They were recruited through convenience sampling, and all either teach courses in which the application could conceivably be used or are closely related to vmbo. Table 1 summarizes the six participants of this study.

*Table 1: Participants of the first preliminary study. [a]Levensbeschouwing, [b]Maatschappijleer. *indicates where teachers have taught or are teaching.*

ID	Gender	Experience (in years)	Course	Level of education*
D1	F	9	Philosophy of life[a]	havo/vwo
D2	F	30	Philosophy of life	vmbo/havo/vwo
D3	M	13	Civics[b]	vmbo/havo/vwo
D4	M	27	Math & Head of department vmbo	vmbo/havo
D5	F	26	Civics	vmbo/havo/vwo
E1	M	25	Educational software & former teacher Dutch language	-

3.1.2 Analysis

Participants were interviewed through Microsoft Teams. After verbal consent was given, the interviews were recorded. The interviews lasted approximately 50 minutes each. In total, five hours and five minutes of audio were recorded. Participants were pseudonymized (see Table 1) and both audio and transcripts were securely stored in Yoda.

Afterwards, the interviews were transcribed. Anneleen Janssen and I both individually marked sentences or phrases within the answers that we deemed important or useful, and then discussed the differences in the sentences we marked. We then used Mural[6] to categorize our findings. In total, we divided 397 sentences or phrases into 14 different categories. Finally, we identified recurring ideas, values, opinions, and other contributions, and merged them into 40 issues that

[6] https://www.mural.co/

shaped the input for the context of use and eventually the requirements. These 40 items can be found in Appendix B.

3.2 Context of use

This section describes the setting in which the anti-filter bubble application will be used. It contains a section on the users that will be using the application, and a section on the contexts that it will be used in.

3.2.1 Users

Students. A key point from previous literature is the broad diversity within the group of teenagers. The teachers also noted this: students differ from each other in preferred learning styles, personalities, interest in certain topics, etcetera. They can also have disabilities such as autism, ADHD, dyslexia, and colour blindness, further diversifying them. Even within the relatively small target group of the current anti-filter bubble app (vmbo-students aged 12-14), there still is a lot of variation between students. However, teachers also mentioned several characteristics specifically inherent to vmbo-students. These students are more easily distracted than students from other levels of education, they want to 'do' much more than read, listen, or watch, they will do less for themselves and more for you as a teacher, and they need more variation in class.

Students do use technology - their phones and laptops - a lot, and mainly play games, use social media such as Instagram or Snapchat and watch YouTube videos. How much the students share about their online world depends on multiple factors; older students are willing to share more, and the closer they are to their teacher the more open they are about their social media usage. However, as noted in previous literature, teenagers are not digital natives as is often assumed. Teachers support that claim. Students are not always technologically capable - save for individual differences - and they cannot be expected to properly work with technology without instruction. They also do not have the skills to handle online information. According to the teachers, they will just use the first Google search result without checking authenticity, or just believe everything they read online to be true. Students' general knowledge of the world has also dropped, and it cannot be assumed that they are aware about topical issues or news events.

Interestingly, whereas related work often mentioned the influence of peers on teenagers, the teachers predominantly noted the influence parents still have in early secondary school years. Parents especially affect the way or the extent to which the teenagers know about current affairs and their personal opinion on the matters.

Teachers. The crucial role of the teacher in a lesson with the filter bubble application was once again underlined by the teachers we spoke to. They are the ones making sure of a safe environment for the students, making sure learning goals are achieved and guiding the learning process of the students.

Teachers disagreed about whether that task could be accomplished by any teacher. Given the topic of filter bubbles and the delicacy that needs to come with discussing polarisation, some

teachers cast doubts about other teachers' technological knowledge needed to answer students' questions, or other teachers' discussion-guiding skills needed for leading the class through delicate topics. Other teachers on the other hand believed that every teacher should be able to handle in class discussions, as that is what they were trained for. Something they did agree on is the wish for a 'guide' for teachers, containing the necessary knowledge and some discussion starters.

3.2.2 Context

Physical context. The application will be used in a classroom. In a traditional setting, this would mean around 25 students seated in pairs, with the teacher's desk at the front next to the smart board. However, this will and can differ between schools. The physical context is therefore not predetermined.

Social context. The importance of a safe environment for students where everyone can say what they want has been established previously, but the degree in which this is actually possible in a class depends on the actual makeup of that class. It differs per class what topics are open for discussion. Some teachers did not foresee any problem, some noticed some opinionated bubbles forming in their classes and some even told us that the current atmosphere in their class was not safe enough to discuss certain sensitive topics.

Teachers stressed the importance of a social bond with students, especially in vmbo. Related to the point that vmbo-students will work more often for you as a teacher than for themselves, a good relationship with students strengthens their work ethic and improves the students' feeling of safety in class.

Organisational context. Different schools have different curriculums, which shows in the opinions of teachers on where the application would fit into their school. The obvious place would be *maatschappijleer* (civics or citizenship education) for some, but some schools do not teach maatschappijleer in the lower years, while some of them do offer other related courses such as lifestyle informatics or information sciences. In general, lessons typically take between 45 minutes to an hour. Other schools have scheduled time slots for weeklong educational projects, commonly reserved for topics like alcohol usage or traffic dangers.

Technological context. Accelerated by the necessity created through the Covid-19 pandemic, a number of schools actively use laptops or tablets in their educational programme. Some call themselves 'Google-schools'. Most laptops used are indeed (Google) Chromebooks, and several teachers mentioned the use of the online Google Classroom environment. Teachers post homework assignments online, students can hand in their homework online, and this use of technology also facilitates communication between teachers and students. The extent to which 'digital' schools make use of this online environment also depends on teachers' preferences.

Digital schools or not, all teachers mentioned some way they make use of technology in their classes. They use the smart boards - present in almost all classrooms - to present interactive quizzes like Kahoot or Mentimeter, or show content-related videos to strengthen their lessons.

3.3 Requirements

Based on the related work and the above context of use, which emanated from the six interviews with teachers and experts in the field, the following six requirements were established.

Req 1. The application should be part of a lesson or project.

Teachers agreed unanimously that one-time usage of the application is not adequate:

> *"One themed lesson is fun, and you'll maybe have their attention for a short while, but I don't think it would last."* (D1)

> "*Working in projects, having multiple days to work with one theme, that would stick a lot better."* (D5)

It is therefore important that the application is part of either a lesson plan, or an educational project akin to alcohol or traffic educational programmes. Teachers had differing opinions and ideas on where the application could be used exactly. This is however also dependent on characteristics of individual schools, such as the courses they are offering, how their lessons are structured, and if there are designated possibilities throughout the year for educational projects. The definitive place of the app in a curriculum should therefore be decided on an individual basis. A related important insight to note is that content should not be built around the application, but the application should be a tool to support (either pre-existing or specifically developed) content. This also prevents the application from becoming a single-use gimmick:

> *"But it* [content] *should not be neglected. As if the app is the main theme, and that there should be something surrounding the app.* [...] *It should be the other way around."* (D4)

> *"It* [the application] *should not be a toy, like a fun gadget to use in class."* (D5)

Req 2. Teachers should have a clear overview of what students are doing within the application.

Regardless of whether students are working individually or in small groups, it can be hard for teachers to track what their students are doing in the application. Time constraints are most frequently mentioned, as lessons are not long enough to talk to every individual or group extensively. Additionally, students can be very protective about showing what they are doing within educational applications:

> *"When students are working within an app, it's not always easy to supervise them. [...] They only call for you when something is not working, that's my experience."* (D1)

It is important that teachers do have an overview of what students are doing for multiple reasons. First of all, they need to be able to maintain a safe environment within the class - see Req. 3. They also are the ones ultimately responsible for making sure all learning goals are reached, so

they need to know how the students are doing. The exact way of implementing such an overview for teachers needs to be subject to further research:

> *"How will a teacher gain insight into what a student has done? How will you capture that? Will it connect to a dashboard? Can you use that for another teaching method? Can you use it on all platforms in the same way?"* (E1)

In order to better grasp the progress of students, teachers also underlined the importance of understanding the learning goals set forward by the application. Teachers need to know what principles the application and its content are based on, what it can add to the classroom and how they can fulfil its potential in their teaching:

> *"There should be a solid guidebook accompanying it, so that it's not like 'here is the app, good luck and have fun', but what the background of the application is, what the goal is, how to properly implement it."* (D5)

> *"You should make clear* [to teachers]*, 'this app will help you with this and that, because it's about this and that. And it even aligns with this core objective or that final attainment level'."* (E1)

Req 3. The application should (help) foster a safe environment for students within the classroom.

When considering prerequisites for discussing filter bubbles and sensitive topics like polarisation in the classroom, all teachers stressed that every single student should feel safe at any point and never be afraid to tell what they believe:

> *"You need to make sure, especially when dealing with sensitive topics, that every opinion can be heard, that it is safe, and that sometimes you can't discuss a topic because it is just too sensitive."* (D3)

> "*Everyone needs to be able to say what they want or what they feel. And in one class, that will be easier than in another*." (D4)

Teachers play a crucial role in maintaining this safe atmosphere. They deem it their responsibility to make sure that every student is able to say what they want to say. The application should be able to support their efforts:

> "[As a teacher] *you need to intervene at the right moments, like 'now you are taking it out of context', or 'you are allowed to think that, but we are talking about something else right now'."* (D1)

> "[If you're discussing sensitive topics] *you need to be very skilled in managing such discussions."* (D2)

> "*That* [creating a safe environment] *is kind of a task for a teacher, or also for the application* [...] *I should feel that I can actually doubt my own perspectives or other's views* [as a student]." (D3)

Teachers provided multiple ways to maintain a safe environment, like "*focussing on knowledge*" (D2) or by "*adding science to a debate, adding the statistics"* (D3). Interestingly, some teachers noted that because their role is so important in discussing sensitive topics, not all teachers are necessarily readily equipped to do that. The application should help those teachers with the topics at hand or at least alert them to the knowledge or skills necessary:

> "*If you're a teacher who does not have the right tools for that* [discussing sensitive topics], *then it would be wiser to say 'I'm not covering this topic'.* [...] *It can do more harm than good if you have the wrong teachers in front of class.* [...] *That doesn't mean that a second-grade math teacher can't do this. But maybe the instructions should be clear. You could make a kind of disclaimer, or an introduction, like 'you should have these skills as a teacher'."* (D3)

> "*Of course, there are a lot of teachers who are very reluctant to discuss very sensitive topics.* [...] *If you provide them with a helping hand, even if it's only a short clip or cards with prepared questions, it'll help to discuss these themes."* (D4)

Req 4. Students should work together with the application. Sharing experiences together and discussing increases curiosity.

When discussing lesson formats with the teachers, they all indicated that students working individually would have less of an impact than students working together in small groups;

> *"I think you should never do that individually* [...] *but that you can compare something with each other. As in, 'hey, we're both doing the same, but we are getting a different answer'. I can imagine that that invokes curiosity."* (D4)

> *"I think it* [using the app] *should not be individual. Definitely in smaller groups. With four people or something like that, they can quickly exchange things and see each other's work. Exchanging things with each other, I think that would be the most effective."* (D2)

Teachers also provided insight in how this working together could come to fruition:

> *"Collaborative learning* [works well], *when there is an element of competition for example, or when they quickly have to produce a product before the end of class.* [...] *Competition works everywhere.* [...] *Just because there is competition, it suddenly just works."* (D1)

Req 5. The application should use meaningful examples and experiences that are relevant for students.

When making students aware of filter bubbles, their consequences, and what they can do about them, it is important to supply enough real-world examples that are meaningful for the students:

"*Especially if you want to do this in younger classes, it should be really close to their perception of the world. And with examples of how it works.*" (D1)

"*I think it's important to make sure that students feel like 'I want to know more about this'. I think you achieve that through using recognizable context, by for example connecting to current events.* [...] *So try to apply it* [the app] *to real contexts. Try to invoke 'oh yeah, I see this in real life as well'.*" (D3)

"If they see a certain necessity or take interest in the topic [...] *the chance of success will be higher. They need to either see the benefit of it, or think it is interesting enough to just like it."* (D4)

Two examples specifically related to filter bubbles were provided by all teachers; the Dutch talk show *Zondag met Lubach* and the Netflix-documentary *The Social Dilemma*. Students seem to be interested in those, watch them with interest, and learn something from them. However, both are mainly about the extremes in society, making them less relatable for students:

"*A movie like The Social Dilemma will definitely help. Because that one is easy-to-understand, and the children understand it immediately because it is their world.*" (D4)

"*The problem with that 'fabeltjesfuik'* [Zondag met Lubach segment] *is that it's about those loonies, and now they* [students] *have to apply that to their own lives. So that's still a step that you have to take.*" (D1)

Several teachers emphasized that it is not just about providing relatable and realistic examples, but that students really have to experience something within the application:

"The good apps with content often really have a very special experience. The experience of that content ensures an app sticks." (E1)

"*Something that really works is that people can experience something. I really believe that, that is very powerful.*" (D3)

Req 6. The application should cater to all students: from those with low or high interest or knowledge in the topics, to those with certain disabilities.

Something that already emerged from previous literature is that students cannot be lumped together into one group with the same needs, preferences, and attitudes. Teachers really stressed that no student is the same and that there are considerable individual differences. These individual differences also manifest themselves in the way they handle (online) information:

"*You can't talk about 'the students' and have one image of them all. That's impossible.*" (D2)

"[How they handle information online] *really differs.* [...] *You can really see the difference between the student who can talk with his parents about what they see and hear, and those who don't.*" (D1)

It is important that the application caters to all students, not just a select group of students based on their knowledge level of filter bubbles or their skills to deal with online information, but also on their interests, preferred learning styles or possible disabilities:

> *"I think it* [using direct or story-based activities] *really depends on the type of student. And the type of teacher.* [...] *I think both can be effective, and I think it's good if you also try them both.* [...] *So tailored help, distinguishing between students.* (D3)

> *"That is the trick. How can you make it appealing, while you are dealing with so many audiences, with so many personalities. Introvert to extrovert, autistic, ADHD, dyslectic* [...] *There are also students who don't know anything about this* [the filter bubble], *and those who already know a lot about it. You'll need to tend to them both.* [...] *So a bit of differentiation in what you are offering would be a good idea.*" (D4)

3.4 Storyboarding

Based on the context of use and the requirements, we developed two storyboards that visualize two possible implementations of the application in class. Both consist of five frames that depict an activity that could be part of the application. They were used as input for pre-study 2, described in Chapter 4. This section describes the contents of the frames of both storyboards.

3.4.1 Storyboard 1

The goal of the activity presented in the first storyboard, *Volgende Video*[7] (Figure 2), is to let students realize that there are differences in what YouTube recommends to everyone. Because of its straightforward nature, it was intended to serve as an introductory activity.

Frame 1. The class has been divided into groups of three. The teacher introduces the activity, and instructs students to get their phones out and navigate to the app. The teacher starts the activity through the smartboard.

Frame 2. On their phones, students can read the instructions of the activity once more. Together with their group, they decide on a video to pick for the activity. In this example they can choose a video on TikTok, Formula 1, Minecraft, or The Voice.

Frame 3. Students have picked the video on The Voice. The video opens in the YouTube application on their phones, and thus shows their personal recommendations for the next video to watch.

Frame 4. Students discuss differences and similarities between their lists of recommended videos. They return to the anti-filter bubble application, which prompts them for an answer to the question 'what did you notice?'.

[7] Next Video

Frame 5. After all groups have submitted an answer, the smartboard in class shows all of the groups' answers. The teacher can start discussing these answers with the class.

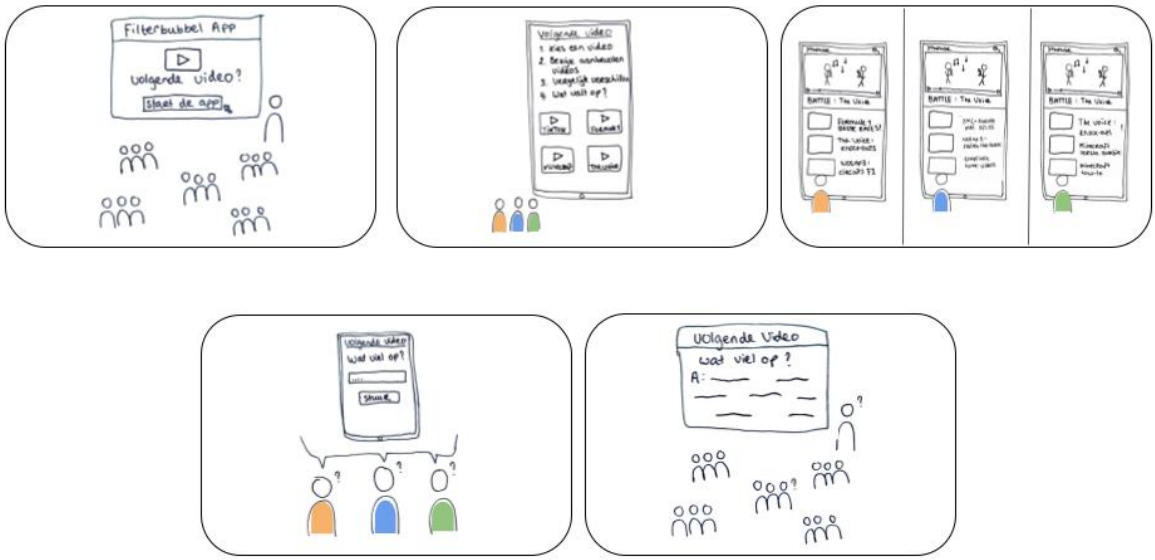

Figure 2: Storyboard 1.

3.4.2 Storyboard 2

The second storyboard (Figure 3) was based on an activity that was devised in the 2020 Computer Science software project (J. Berkhout, personal communication, 2020), then called *Aanradend Algoritme*[8]. In this activity, students take on the role of an algorithm and recommend YouTube videos to a user to keep them on YouTube. The goal is to make students aware of the inner workings of the YouTube algorithm and recommendation algorithms in general. The activity as described in the software project was adapted slightly in order to better fit the current requirements and context of use, such as the benefit of working in groups.

Frame 1. The activity starts with the introduction of a viewer engaged in a video on YouTube. In this case, that video is about a red square. Students are told they will adopt the role of YouTube's algorithm and have to decide what the viewer will watch next, with the goal of keeping them interested.

Frame 2. Having been divided into groups of three, students can decide on their phone what the viewer should watch next. In this case, they can choose between a video on two red squares or on a yellow triangle.

Frame 3. Students discuss the choice they will make as a group.

[8] Recommending Algorithm

Frame 4. After making a decision, they will see the impact of their choice on the smartboard in front of class. The interest of the viewer either rises or decreases, depending on the choice they made. In this case, they chose the yellow triangle: the interest of the viewer drops.

Frame 5. At the end of the activity, an overview appears on screen comparing how well all groups maintained the interest of the viewer, adding a layer of competition to the activity. The teacher can then start a discussion about why some groups did not manage to maintain interest of the viewer while others did.

Figure 3: Storyboard 2.

4. PRE-STUDY 2: STORYBOARDING

To further gain an insight into what would be feasible and effective activities for the anti-filter bubble application, a second preliminary study was done using the storyboards from Section 3.4. This chapter reports on that study and concludes with the issues and topics that need to be considered when developing further activities and the first lo-fi prototypes. Section 1 describes the methodology, participants and methods of analysis, Section 2 presents the results from the interviews and Section 3 contains the interpretations and consequences that these results have for the next stages of the project.

4.1 Methodology

We carried out three semi-structured interviews with teachers in secondary schools. In these interviews, we discussed their use of technology in class and their students' online behaviour, similar to the previous pre-study. However, the main goal of these three interviews was to use the storyboards to provoke the teachers into telling us what would work and what would not. The protocol for these interviews can be found in Appendix C.

4.1.1 Participants

Three teachers took part in the interviews with storyboards, all recruited through convenience sampling. Equally to the participants in pre-study 1, all three teach related courses where the anti-filter bubble application could conceivably be used. Table 2 summarizes the three participants of this study.

Table 2: Participants of the second preliminary study on storyboards. [b]Maatschappijleer.

ID	Gender	Experience (in years)	Course	Level of education
D6	F	3	Civics[b]	vmbo/havo
D7	M	16	Computer Science, Biology & Head of department havo 4/5	vmbo/havo/vwo
D8	M	6	Computer Science & History	havo

4.1.2 Analysis

All interviews took place through Microsoft Teams. After verbal consent, the interviews were recorded. The storyboards were shown using slides in a presentation, with a final slide showing all five frames to provide an overview to participants. The interviews lasted approximately 55 minutes each: in total, two hours and 45 minutes of audio were recorded. Participants were pseudonymized (see Table 2) and both audio and transcripts were securely stored in Yoda.

Similarly to the interviews of pre-study 1, the interviews were transcribed, individually marked and then discussed. Mural was used to categorize findings. We marked a total of 192 sentences or phrases which we divided into 18 categories, 13 of which were also used in pre-study 1, with the other five new categories being specifically about the storyboards. From these 192 excerpts,

we extracted 29 new insights. Eight of those were attributed to the previously used 13 categories, the other 21 insights were related to the storyboards. The full list of insights can be found in Appendix D.

4.2 Results

The 29 insights were either related to the context of use, storyboard 1 or storyboard 2. They are presented here as results complemented with direct quotes from the participating teachers. The next section contains the implications for the next phase of development.

4.2.1 Context of use

Users. Teachers restated much of the same points found in pre-study 1. Students are dissimilar in their (online) behaviour, their knowledge of the topic, and their interests. Another point mentioned again was the need for vmbo-students to actually work on something and the need for extrinsic motivation. One teacher disagreed with this and posited that this might actually be a case of the Pygmalion effect, where low(er) expectations can lead to low(er) performances.

Regarding the teachers themselves, two topics emerged that were not as prevalent in pre-study 1. First, they said that providing the students with predetermined learning goals and a fixed lesson structure is important and benefits students. Second, aside from being able to handle discussions on sensitive topics, teachers should also have sufficient knowledge on the actual topic at hand - filter bubbles and personalisation algorithms. Once again, they all suggested supplying a 'manual' or at least guidelines with the application, on both a knowledge level (what they need to know before tackling the topic) and a meta level (what they need to be able to do before tackling the topic).

Context. Again, teachers had different ideas for the place of the application in their school's curriculum, similarly to the previous pre-study. They did stress the need for plenty of room for discussion in their lessons, both before and after the activities with the application, as they expected their students to raise questions:

> *"I think that when they're done with this, they'll have a lot of questions afterwards, like 'how can this be and what happened there'? And those questions will only arise after they have experienced it."* (D6)

> *"I think the app is really good for that bit of realization, as in 'hey, something weird is going on here, this is apparently happening. And then the teacher for the part of 'but what do we think about that? And should we think something of that?'* (D7)

The importance of a safe environment for students was also stressed once more, underlining the universality of that precaution among teachers. There is also a major role for the teacher in managing this:

> *"Without proper guidance it can be unsafe. Under good circumstances, it can actually make for a very nice dialogue, but the circumstances have to be right."* (D7)

> "*Especially in vmbo, it is so important to have a strong bond with your class before you start working on such* [sensitive] *topics. Because that safe atmosphere has to be there, otherwise it won't work."* (D6)

4.2.2 Storyboard 1

Teachers commented positively on the collaboration aspect of this activity. Not only can students discuss their results within their own group, but the subsequent class discussion can also be held with groups instead of individual students. They also expected students to like the activity, even if it is solely because they would be allowed to use their phone:

> *"I think so [students will like the activity]. Just because of the fact that they are using their phone. They always see that as a plus."* (D6)

All three teachers stressed the need for further, strong discussions after the activity. It is there where students should learn and realize what could be the cause of a difference in recommended videos. This was also related to the worries that teachers had about the application's prompt given to the students ("*what did you notice"*), which they felt was not specific and tangible enough to generate thoughtful answers.

> *"I think that as a teacher you will primarily have a role in that last part, so in that bit of 'what did you notice'. And then make a connection with 'how is that possible'.* (D6)

> "*I think literally the last thing you say, 'what is the effect of this', that is the most important step* [...] *Because that dialogue is not in the app itself. And then it will really depend on how a teacher guides that whether or not it will be beneficial."* (D7)

D8 raised the point of anonymity. The activity heavily relies on students using their own YouTube-account to compare with others, but students might not be willing to share the videos in their recommended sections with everyone in their class.

> *"It really depends, but for some children it is confronting* [to share their recommended videos]. [...] *But that also depends on what the atmosphere is like in the class, if it is safe, how the teacher handles it, in that respect."* (D8)

4.2.3 Storyboard 2

Storyboard 2 elicited more pronounced positive responses from the teachers in comparison with storyboard 1. Teachers expected this activity to raise awareness even before the in-class discussion:

> *"If I had to say right now 'storyboard 1 or storyboard 2', then I'm more enthusiastic about storyboard 2, because I think that the element of surprise is much greater here."* (D7)

> *"In principle, this would be an extremely good approach to create awareness of that* [the filter bubble on YouTube]" (D8)

D6 added that students would need intermediary feedback to get the most out of this activity:

> *"I think that* [intermediary scores/feedback] *is very important for creating awareness. I think it is important that at some point in time, they see: what would have happened had I made another choice?"* (D6)

In this storyboard, topics of videos were indicated with red squares and yellow triangles, as a means to conceptually convey the idea of recommending the right videos. When asked what topics could take their place, two teachers indicated that in order to make the activity more in-depth, actual polarizing or conspiracy-themed videos could be used. As D7 said, "*because you want to show that you will end up in a weird corner* [of the internet]". However, they also said that *"you do not want them to see very extreme videos that are not appropriate for them"* (D7).

When asked about the possibility of collaboration for this activity, teachers responded differently from each other. One teacher explicitly mentioned that this activity would work better when performed individually, while one other said that the possibility of discussing in small groups when picking videos together would be best.

4.3 Discussion

Based on the results above, these are the implications for the further development of these activities and possible future activities.

Safety and teachers' competence. Much like the teachers in pre-study 1, the three participating teachers in this study stressed the importance of a safe atmosphere in class. By now, it is also clear that there is no standard recipe for a good environment like that. Different classes respond differently, different teachers have different teaching styles and competences, and different schools have different guidelines. Additionally, especially given the unclarity surrounding the place of this application within the curriculum, it is by no means certain that the teachers in question will have the necessary experience for guiding class discussions on sensitive topics. The participating teachers in the two pre-studies were also worried that not all teachers might have adequate knowledge on the topics at hand. They suggested a 'manual' might help those in need of more background information, questions, and points of discussion, but worries about the safe atmosphere remain. Therefore, in addition to the manual, a possibility could be to offer choices to the teacher in terms of topics. Certain topics might not be suitable to cover in certain classes, while other classes could go more in-depth on actual real world polarizing topics.

Storyboard 1. The participating teachers showed more interest in Storyboard 2 than in Storyboard 1, and generally reacted more positively to the second one. This appears to be related to two qualities of Storyboard 1. First, it is a relatively short activity, with little to no role for the application itself - students have to go to their own YouTube page, discuss the differences among themselves, and the application 'only' serves as a framework for the entire interaction. The added value of the application is therefore limited in this case. Second, as students have to use their own YouTube apps and accounts, students and teachers might take issue with the potential privacy infringements. Students might not want to share with their entire class what YouTube

recommends to them, and teachers certainly do not want to force students to do it anyway. While a possible solution could be to provide (some) students with fake YouTube profiles and recommendations, that takes away the 'realness' of the entire activity, which was primarily the point: show students that this happens in their own lives too and is not just hypothetical. Another solution for this privacy issue therefore has to be found. The potential in this activity does lie in the aspect of collaboration and comparing with others, which teachers said would be both fun and educational.

Class discussion. It is clear from the first two pre-studies that all teachers expect the class discussion after the activity to bear the educational value of the entire experience. The activities will not inherently be *the* learning experience. This is important to realize, as that means that in terms of duration the activity should probably not take any longer than 10 minutes. It also means that the role of the teacher with this application is even more significant than previously considered.

Collaboration. Interestingly, teachers disagreed on whether the second activity (Storyboard 2) would be better suited for group work or individual usage. Requirement 4, based on pre-study 1, posited that *'Students should work together with the application. Sharing experiences together and discussing increases curiosity'.* While the second part of that requirement is still applicable, the first might not be set in stone. Related to the previous point of the value of a class discussion: if students benefit from doing the activity individually and then discuss their experiences in class, then they should not be forced to do group work in the activities. Either way, the advantages, disadvantages, and other secondary aspects of the issue of collaboration should be further investigated throughout the design phase of this project.

The next step, after these two first pre-studies, is to start developing low-fidelity prototypes. These prototypes will then be evaluated with students and teachers, and iteratively developed based on their feedback. This process is described in the next chapter.

5. DESIGN

This chapter describes the entire design process for the activities that were developed for the anti-filter bubble application. The chapter has been split up in sections describing each activity separately, preceded by a method section that describes the interviews and focus groups that have taken place to gather input on the prototypes. Figure 4 visualizes the design process for the activities and an overall idea for the application.

All prototypes as shown were built in Figma[9], with the exception of one of the final versions of the prototype of Activity 3, which was exported from Figma into ProtoPie[10] to make use of interaction design features not available in Figma.

5.1 Methods

During the design phase of this thesis, both students and teachers have given their input on the conceptual ideas and prototypes that we presented to them during focus groups and interviews. This section describes the setup, participants, and method of analysis for both the interviews with teachers and focus groups with students.

Whereas the pre-studies' participants were solely participating in the context of this thesis, in this case, the pedagogy students also involved in the anti-filter bubble project helped and participated in some of the focus groups and interviews. The degree of cooperation varied from transcribing parts of the audio to conducting interviews together. Some focus groups with students were conducted only by the two pedagogy students. Their questions and discussions emphasized students' online behaviour and their knowledge on algorithms and filter bubbles, but this information was also beneficial and even necessary for some of the design decisions. For more information on their work, I refer to the theses of Zena Bani and Anouk Adriani.

Similarly to the pre-studies, participants were pseudonymized (Dx for teachers, Lx for students) in the transcripts, and all audio and transcripts were stored in the University's secure data deposit Yoda[11]. The number of participants (both students and teachers) we talked to for each activity is shown in Table 3.

[9] https://www.figma.com
[10] https://www.protopie.io
[11] https://www.uu.nl/en/research/yoda

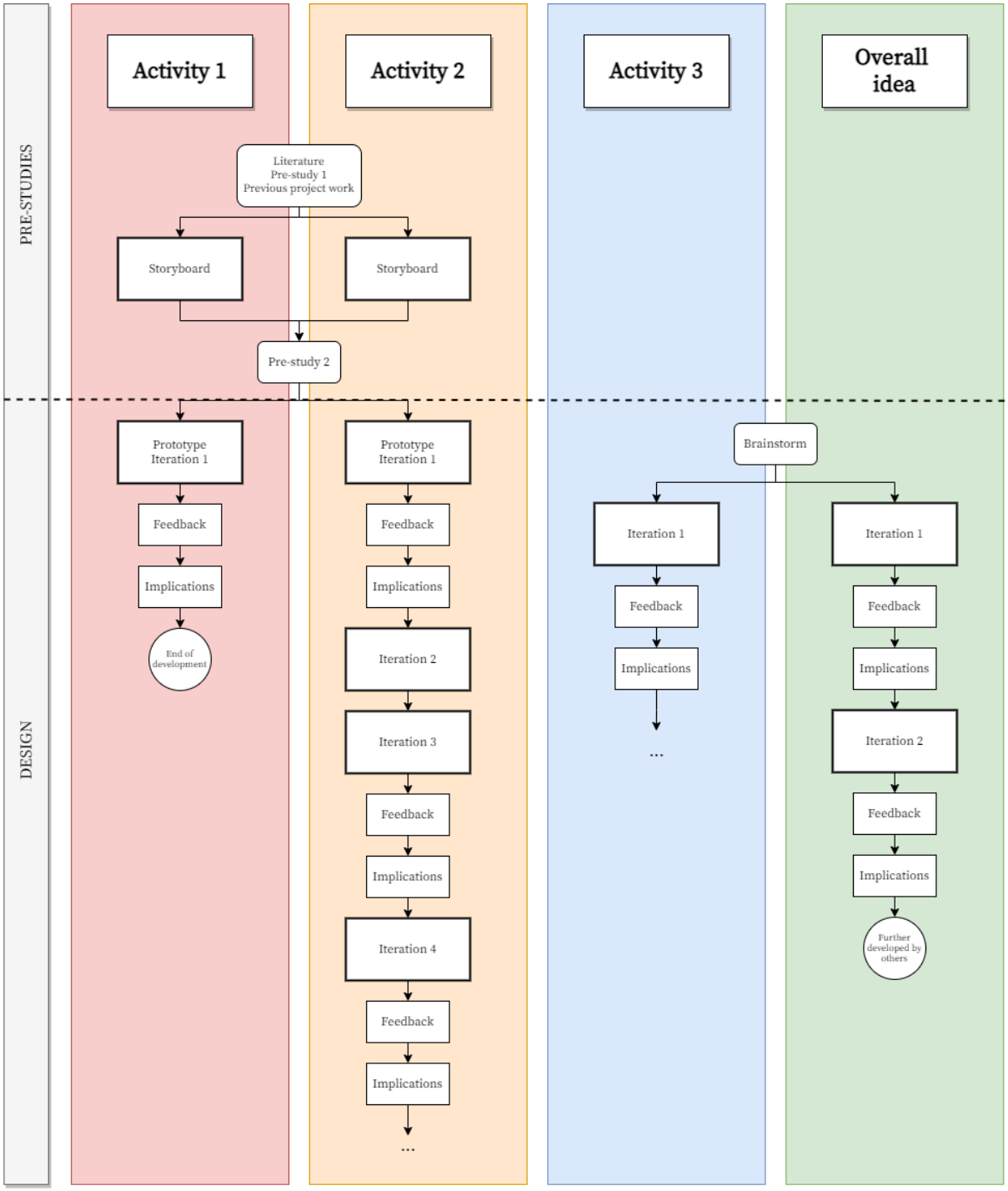

Figure 4: Visualisation of the design process of the three activities and the overall idea. All rectangles correspond with a (sub)section in this design chapter.

Table 3: The number of participants that we showed each prototype to in the design phase.

Prototype	Students	Teachers
Activity 1	4 (pilot)	2
Activity 2	4 (pilot) + 15	6
Activity 3	5	3
Structure	-	4

5.1.1 Interviews with teachers

Contrary to earlier interviews with teachers, these interviews were not aimed at gaining insight into the context of use of the application and the learning processes of students, but to directly garner feedback on the prototypes. The goal was threefold: find out if teachers thought their students would like the activity, if the activity could conceivably be used in one of their classes, and how they envisioned their role (as the teacher) within and during the activity.

5.1.1.1 Method

We conducted semi-structured interviews with participants, which allowed us to talk the participant through the prototype of the application (i.e. what a student would experience doing the activity). That walkthrough prompted the teacher to tell us what they liked about the activity, what they did not like or would like to see improved. We then asked questions about specific elements of the activity, such as the difficulty, and about elements related to the integration of the activity into the classroom, such as collaboration and duration. All protocols with the predetermined questions and topics can be found in Appendix F.

All interviews took place through Microsoft Teams. The interviews were recorded (audio only) after verbal consent was given.

5.1.1.2 Participants

Throughout the entire design process, we spoke to seven teachers, one of which we interviewed two times. D9 & D10 were interviewed together, as were D12 & D13, and D13 once more with D14 and D15. Table 4 provides an overview of the participants of this study.

Table 4: Participating teachers in the design phase. [b]Maatschappijleer.

ID	Gender	Experience (in years)	Course	Level of education
D9	F	13	Economics & tutor in vmbo	vmbo/havo
D10	M	12	Geography	havo/vwo
D11	F	3	Civics[b]	vmbo
D12	F	7	Civics[b]	vmbo/havo/vwo
D13	F	14	Civics[b]	vmbo
D14	M	21	Civics[b] & History	vmbo
D15	F	1	History (trainee)	vmbo

5.1.1.3 Analysis

A similar method of analysis to that of the pre-studies was used for these interviews. The total amount of audio recorded was three hours and 25 minutes. This audio material was transcribed and once again, both Anneleen Janssen and I individually marked interesting and insightful sentences or phrases. Mural was used to categorize the findings. From these interviews, we extracted 236 sentences or phrases, and placed them in 20 different categories - most sentences or phrases were related to the categories belonging to the prototypes, but some utterances provided insight in for example their students' online behaviour as well. Throughout the sections on the design processes of the activities, insights and comments from teachers will be brought forward when relevant.

5.1.2 Focus groups with students

The most important feedback on the activities undoubtedly comes from those that will eventually be using the anti-filter bubble application: students. As mentioned before, the current focus of the project is first or second grade (12 to 14-year-olds) vmbo students.

5.1.2.1 Method

To get the most out of the sessions with students, we opted for conducting focus groups. Focus groups allow students to react to one another and makes them more at ease and comfortable with sharing their opinions and ideas.

Most focus groups were set up to first play the activity with the students, talk and guide them through it, and then talk with them about their experiences. Questions about their experience were often interwoven with questions about their own social media or internet usage, as the activities sometimes provided the perfect link to discuss that. Primarily however, questions about the activity itself were asked. These included direct questions about their thoughts (e.g. 'would you like doing this in class?'), meta-questions about their reasons for certain interactions (e.g. 'why did you choose this video?'), and questions to prompt them for suggestions and ideas

about what they would like to see in the activity or application. All protocols with questions and topics can be found in Appendix G.

One focus group deviated from this general setup. Instead of reacting, responding, or thinking about one specific activity, we challenged three students in a brainstorm session to come up with ideas for activities within the application. Focus groups lasted 30-60 minutes (depending on the situation), and almost all focus groups took place in a physical setting in a classroom at the students' schools. One school (in total six participants) preferred talks to be held online because of Covid-19. Covid-19 also made it impracticable for us to offer any compensatory refreshments to the students.

Beforehand, one pilot focus group was held at an Amsterdam secondary school with four students. This focus group was not recorded, as these students were not officially part of the anti-filter bubble project and their informed consent did not include audio recording. The pilot helped to iron out practical flaws of the protocol and gather some first feedback on the prototypes.

5.1.2.2 Participants

Excluding the pilot focus group, 20 students aged 12-14 participated in the design phase of this thesis (12 female, 8 male). Three of those students (L1, L2 & L3) participated twice. Full details can be found in Table 15 in Appendix E.

10 participants were recruited from the schools in Utrecht involved in the anti-filter bubble project. Teachers selected students they thought were able to communicate their thoughts and opinions adequately. This had the added advantage of students knowing each other in advance, allowing them to loosen up more compared to being in a group of random students. The other 10 student participants attend schools in Amsterdam. They are in the so-called *Kopklas,* a bridge year between the last year of primary school and the first year of secondary school. Although these students sometimes struggle with a deficit in the Dutch language, they are also better trained in giving feedback to each other, making them more than qualified to provide input on ideas and prototypes.

The procedure was accepted by Utrecht University's ethics committee and informed (parental) consent was gathered and provided beforehand. Students were 'cleared' for multiple moments of data gathering, which meant that when we talked to students, it often was not the first time these students had been questioned. The benefit of that is that students got more accustomed to the setting of focus groups.

5.1.2.3 Analysis

During focus groups, audio was recorded - video recording was not included in the ethics committee's approval. In total, excluding the focus groups conducted solely by the two pedagogy students, 4 hours and 24 minutes of audio was recorded. This audio was transcribed and imported into NVivo[12].

[12] https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home

In NVivo, I individually coded the transcripts using open (hierarchical) coding. Given that the primary goals of these focus groups were to gather information on students' thoughts and experiences with filter bubbles and social media, and to gather feedback, suggestions and ideas on the prototypes, it was not necessary to do any further axial or selective coding.

5.2 Activity 1

After evaluating the storyboard for Activity 1, dubbed *Volgende Video* in that phase, it was developed into a lo-fi prototype using Figma. In response to the feedback given by teachers in pre-study 2, the single open question in the storyboard was replaced by three multiple-choice questions and two open questions to produce more valuable answers. Aside from usability-related issues, nothing else was changed in terms of interaction or visible screens.

5.2.1 Prototype

Below, Figure 5 shows five screens of the lo-fi prototype.

Figure 5: Five screens of the first prototype of Activity 1.

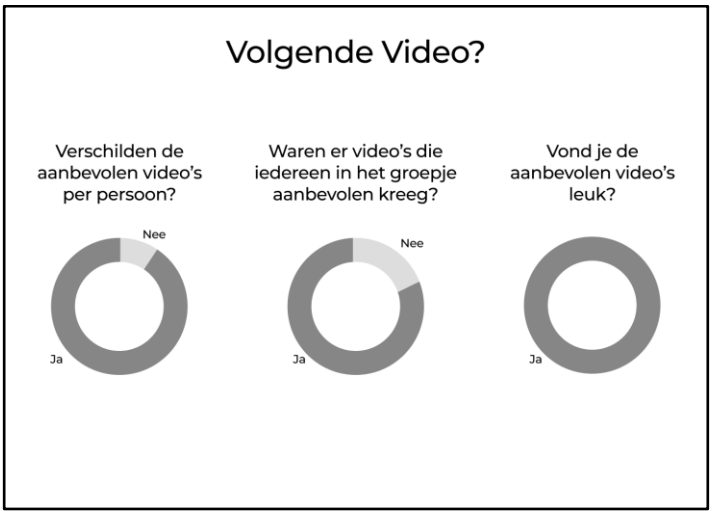

Figure 6: results of the multiple-choice questions are visible on the smartboard in class.

5.2.2 Feedback

The activity was shown to D9 and D10, the students in the pilot focus group, and to members of the anti-filter bubble project team. This section explores the feedback they gave and their implications.

Students. It is important to realize that since this was a pilot focus group, gathering feedback was not the main purpose - rather, it was meant to identify practical issues and flaws in the protocol. However, the feedback the pilot students gave after playing the activity should not be disregarded. Two main issues with the activity surfaced while students were playing. First of all, after clicking through to YouTube, the pilot students became confused about what to do. The application provided instructions beforehand, but that appeared to be insufficient. Second, returning to the application after being on YouTube proved to be difficult for the pilot students. YouTube is 'outside' the application, and students did not understand how or when to switch back to the application. There was the additional problem of some students not having a YouTube-account, and therefore, not having any personally recommended videos.

The pilot students did not have any strong feelings or opinions about the activity, and generally, the activity was met with a sense of apathy.

Teachers. D9 and D10 were cautiously positive about the activity. The changes made to the post-activity questions in the application were well received, as D9 commented:

> *"These are easy questions to fill out, because if you had only put down open questions, you would get like a single answer or something, and that's not the case now."* (D9)

The worry of D8 in pre-study 2 about privacy being an issue for this activity was reiterated by D10. He mentioned that answers given through the application should be anonymous, and that only then it would be okay to play this in class, and not to make it too individual:

> *"You shouldn't be talking too much about personal details* [within the app]. *Then they* [students] *often clam up"* (D10)

Additionally, D9 brought up the role of the teacher, which is rather significant for this activity:

> *"It really is dependent on the teacher how successful this will be. What you offer is great, but how will the follow-up discussion go?* [...] *Everything surrounding it looks good. But the interpretation, yeah, that's pretty teacher-dependent."* (D9)

Project members. The activity was also presented to a number of members of the anti-filter bubble project, primarily professors, researchers and students at Pedagogy. Although they were positive about the concept, their primary concern was the role of the application in the activity. As of right now for this activity, the application is solely used to provide students with instructions, a choice in YouTube-videos, and as a platform for students to answer questions. They argued this could be achieved through verbal instruction and a supplementary Kahoot or Mentimeter as well.

5.2.3 Implications

During these first round(s) of feedback, two main issues emerged aside from practical matters.

First, the element of privacy remains an obstacle for this activity. As mentioned in pre-study 2, several other possibilities such as playing with fake YouTube-profiles were explored, but those would remove the central aspect of experiencing the difference in personally recommended videos with classmates. Second, the added value of the application itself in this activity is negligible. It does not offer more than a traditional verbal instruction combined with a digital tool that is already used in most classes, such as Kahoot or Mentimeter. The role of the teacher in this activity is rather substantial regardless, potentially causing teachers to disregard the application and use their own methods instead.

These (in our view insurmountable) challenges, especially combined with the comparison of how the idea of Activity 2 was received among the same participants, lead to the insight that even if the practical issues would be resolved, this activity would not fulfil the full potential of the anti-filter bubble application. Its further development was therefore put on hold, and after Activity 3 was conceived, it was abandoned altogether. That does not imply that the activity and its evaluations were fruitless; the feedback on for example privacy, collaboration and the role of the teacher are valuable for the development of other activities and aspects of the anti-filter bubble application as well.

5.3 Activity 2

Activity 2, with *Algorithm Activity* and *Aanradend Algoritme* as working titles, was met with enthusiasm from teachers in pre-study 2. Like Activity 1, it was translated from its storyboard into a low-fidelity prototype using Figma. Contrary to Activity 1 however, it underwent several changes when digitized (detailed below). This section describes the continuous evolution of Activity 2 throughout the design phase, broken down in four distinguishable iterations of the prototype.

5.3.1 First iteration

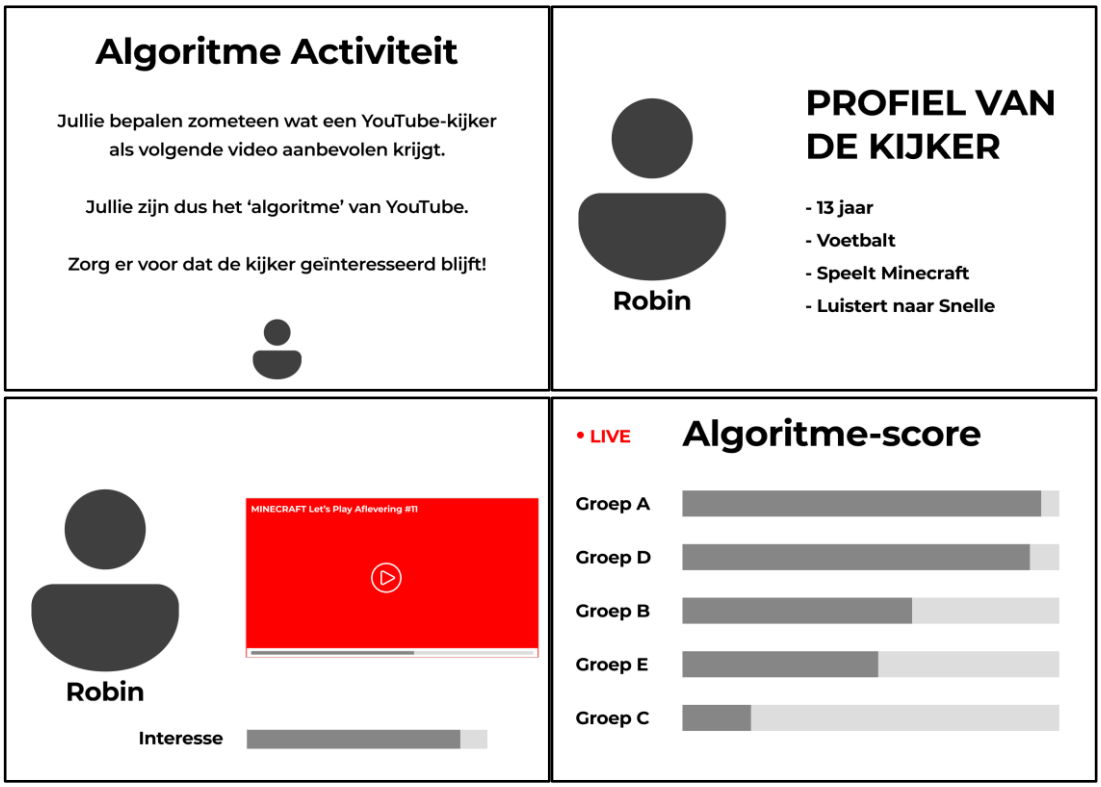

Figure 7: Screens 1 to 4 (left to right, top to bottom).

The activity as envisioned in the first iteration of the low-fidelity prototype is as follows. The first part of the activity is set on the smartboard, projector, or any big screen in class. Students are first explained what they will be doing (Screen 1), after they will get an introduction to the viewer (Screen 2), with a name, age, and a number of hobbies. This profile was included so students would have the ability to empathize with the viewer and picture them as a real human being. Additionally, students are also shown what Robin, the viewer, is currently watching on YouTube, along with an interest meter (Screen 3). After this introduction, students switch to their mobile phones, where they once more have the possibility to view Robin's profile and the video currently being watched.

Then, students need to decide what to recommend to Robin. They have two options to choose from (Screen 4). After picking a video, they get to an intermediary 'confirmation' screen (Screen

5), providing students with some preliminary insight into whether they have made the right choice. Finally, the result screen (Screen 6) shows the effect of their choice on Robin's interest in YouTube and provides feedback on why Robin liked or did not like the video. Teachers in pre-study 2 mentioned that providing this feedback would be necessary to have students understand what the impact of their choice is. While students are playing the activity, the smartboard in class also shows every group's live score (Screen 7). This first prototype contained two complete rounds of play.

Figure 8: Screens 5-7.

The triangles and squares previously shown in the storyboard were replaced with real YouTube-video topics. They are not polarizing, sensitive or difficult, but rather videos that students themselves would watch on YouTube (such as make-up tutorials or Minecraft gameplay videos from prolific YouTubers). This decision was made mainly because this was expected to grasp students' attention more than videos on heavy subjects, not just because it would be more interesting to them, but also more relatable. The videos we picked are real videos by real YouTubers, but we did not include thumbnails for the sake of simplicity of the prototype.

5.3.1.1 Feedback

This version of the activity was shown to the pilot focus group with students, one actual focus group (FG1), and to three teachers (D9, D10, D11).

Students. Both focus groups of students provided a lot of valuable feedback. The pilot students wanted a more comprehensive profile of Robin beforehand, with for example information about what kind of videos he watches. The choices themselves were too easy for them. Suggestions to make deciding more difficult included having more available options (e.g. four instead of two)

and more rounds of play. The pilot students also criticized the video topics for being boring, but also recognized that that is a personal feeling and that some classmates would like these videos. They did mention they would like to see thumbnails for the videos, instead of just the title. Finally, the pilot students appreciated the element of competition in the activity, and added that they would like to see a prize for the group or student that performed the best.

The other focus group (FG1) liked and appreciated the interactivity, but also indicated that the activity in its current form was too easy. To increase difficulty, they suggested fleshing out Robin a bit more so they would have to consider more factors when picking a video. Similar to the pilot students, these students also suggested adding more videos to choose from, and more similar videos, *"so you have to figure out what game* [-video] *would suit him best"* (L2). Students also appreciated the intermediary feedback screens:

> "*I think it gives a bit of clarification, because otherwise it would be like 'you chose this', but why did you actually choose that, why would you choose that, so I think it is a good addition."* (L1)

When asked if they would rather play this activity individually or in groups, one student gave a compelling reason for collaboration:

> *"If you're in a team, you'll have to talk to each other about it, while if you do it individually and you don't really feel like it, you'd be more likely to start guessing."* (L2)

Students also liked the competition displayed through the live scores screen, and related it to Kahoot. Regarding duration, students expected the activity to last between 5 and 10 minutes, depending on the difficulty.

Teachers. D9, D10 and D11 responded very positively and enthusiastic to the activity prototype. They enjoyed the format and genuinely expected students to like it as well.

Regarding collaboration, D9 and D10 both thought this activity would work best if performed individually. Not only would that result in much more variation in reasoning and chosen videos, but they also saw a significant disadvantage of group work:

> *"Group work is always nice, students like it as well. But the problem with groups is: you'll have one draft horse doing all the work, two people sitting on top of the horse, and one being pulled by the horse."* (D10)

However, they also recognized that other teachers might feel differently about collaborating, as their preferences for either working individually or working in groups are *"very dependent on the kind of teacher"* (D11).

On the topics of the videos, teachers agreed with our decision to not use controversial or sensitive videos, but videos that resonate well with the students. Making it as closely related to them as possible would surely improve the students' experience. However, to deepen their

knowledge, D9 also suggested maybe adding the possibility of a second round that does include more controversial videos. D10 and D11 both saw potential in that idea as well:

> *"I think that those* [political/sensitive videos] *would be less interesting, but at the same time I think that's a good thing. Because then they'd get into contact with that and see that there's a difference there as well.* [...] *but then I would rather see that as a kind of sequel or something*." (D11)

In terms of difficulty, teachers estimated that the current level would be ideal for vmbo-*basis* and vmbo-*kader*, but that the slightly more difficult track vmbo-*theoretische leerweg* would need more of a challenge. They expected that the intermediary feedback would not be read by all students, but for those having difficulty with the activity, it would provide just that extra bit of explanation they needed.

Similarly to the students, D11 preferred the activity to last between 5 and 10 minutes. Finally, she also emphasized the importance of the application looking good to engage students.

5.3.1.2 Implications

Conclusively, the first prototype of Activity 2 was received very favourably among both teachers and students. Some matters would need resolving, such as the lack of difficulty or challenge for students, but overall, this first feedback proved promising for the further development of the activity.

As mentioned, the difficulty needs to be raised in further iterations of the prototype. Teachers and primarily students suggested means to do this, of which at least increasing choice and further elaborating on Robin seem most viable. The notion of collaboration deserves attention as well, as teachers and students seemed to disagree on that, which also emerged from pre-study 2.

5.3.2 Second iteration

Figure 9: Screens 8 and 9.

Several changes were made to the second iteration of the prototype. Most notably, we heeded the suggestions of students and increased the number of possible choices from two to four per round. In addition, we also increased the number of rounds in the prototype from two to four

rounds. Both these modifications were aimed at raising the difficulty of the activity. To somewhat prevent students having to pick randomly each round, a 'storyline' was implemented. The first video Robin watches (before the activity) is a Minecraft video. In the first round, the correct choice would be Milan Knol's[13] Minecraft video, after which round two contains another Milan Knol Minecraft video, this time among other videos on games. Round three then contains no game videos, but does display a video originating from Milan Knol's channel. The underlying story here is that Robin, through watching his Minecraft videos, started liking Knol himself. The fourth round further solidifies that with the correct option there being a video from someone else with Knol appearing as a guest. Robin's profile was also enhanced, which was now showing the YouTube-channels he is subscribed to, which is rather relevant information for an algorithm.

Doubling the number of videos and rounds increased the total number of videos to choose from from 4 to 16. We therefore changed the feedback texts to be independent of previously made choices, as making those texts dependent on previous choices with Figma's limited tools for dependencies would have meant creating 256 separate screens. That also meant removing the 'score bar' for now, which we replaced with arrows indicating whether Robin's interest rose or fell by how much.

Visually, a switch was made from mobile screen size (375x812) to desktop screen size (1440x1024), visible in Screens 8 and 9. This was done because the application in its final form would be a web-application, and given most schools work with laptops or tablets, this resolution (or at least a scaled-up resolution) would be primarily used.

This iteration of the prototype turned out to be an intermediary step in the development. After also consulting with other members of the anti-filter bubble project, we decided a number of improvements had to be implemented before further evaluation sessions took place. These are detailed in the next section.

5.3.3 Third iteration

The primary change in this iteration of the prototype was the addition of actual thumbnails. While evaluating prototypes should be primarily about the interaction and flow through the activity and not about content, thumbnails were deemed such an essential aspect that they could not be omitted. They were also already requested by for example the pilot students, and other project members also suggested students could engage more with actual visual thumbnails. Regarding the videos themselves, some minor replacements were made to make the topics more relevant to the students.

Another change made was to the (intermediary) feedback that students receive when having picked a video. This previously included Robin's personal opinion on topics and what Robin would rather have seen. We deemed this was too dissimilar to real-life algorithms, and consequently replaced Robin's thoughts and attitudes with their actual (inter)actions within YouTube, such as liking or disliking a video or sharing it with his friends (see Screen 13).

[13] A popular Dutch YouTuber mainly focused on playing games.

Finally, to facilitate students in their reasoning, we added an 'algorithm table' (Figure 11) for the students to use during the focus groups. This algorithm table provided students with a means to note and write down what they - being the algorithm - figured Robin does and does not like. Due to Figma user input constraints, the algorithm table was implemented physically on paper during the focus groups. In further development stages, the table might be integrated in the application itself.

Figure 10: Screens 10 to 13 (left to right, top to bottom).

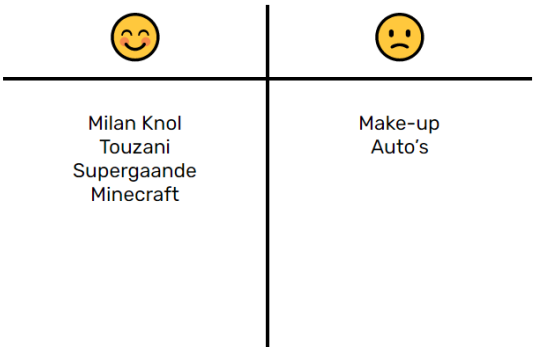

Figure 11: An impression of the algorithm table.

5.3.3.1 Feedback

This iteration of the prototype was used in the second focus group (FG2) and also informally presented to Mira Media, experts on the topic of media literacy in classrooms.

Students. Students from FG2 did not play this activity on their own. They did make the choices themselves and had to provide their own reasoning, but with us guiding them through the process.

They indicated that they did like the activity and would prefer it over their teacher just explaining it to them ("*because you can do it yourself"*, L4). After playing the activity, one student also indicated that they understood the concept of algorithms a lot better now compared to before. The videos and their topics were familiar to all students, although they also indicated they do not necessarily watch these kinds of videos. Despite the adaptations based on previous feedback, students indicated the activity was too easy still. Similar to their predecessors, students did have valuable suggestions to increase difficulty:

> *"I think that first video, in the first round, you immediately put Minecraft with Milan Knol. If you hadn't done that, it would have been a bit more difficult. But even then, I think it wouldn't be that hard, because you can immediately see 'oh he likes Milan Knol, he is subscribed to him, and he plays Minecraft'."* (L6)

All students thought this activity ought to be more fun in groups than individually, especially if groups were able to compete against one another (*"winning is always fun"*, L6). The comparison with Kahoot was once again quickly drawn. They expected that their classmates and themselves would also put more effort into the activity if competition was involved:

> "*I think so* [competition would make it more fun]*, because then you're really going to try a little harder. You'd really play against other people."* (L7)

When explaining their thought processes for picking certain videos, it appeared students really thought through their decisions as the algorithm also would have done. They thought about what kind of person Robin could be, and what kind of conclusion an algorithm would draw from that:

> *"Ultimately, he* [Robin] *is a gamer, and someone who plays Minecraft, well… You know, I don't think that someone like that would like GTA*." (L6)

This kind of thinking was also applied when students were using the algorithm table, which was well-used throughout the session:

> *"Maybe he* [Robin] *likes* [to watch] *challenges. Because he* [Milan Knol] *is doing a challenge to survive 24 hours in a zoo. Maybe that could mean something for the algorithm."* (L5)

Mira Media. Even though showing the prototypes to the two experts at Mira Media was primarily intended to show our progress, the ensuing discussions and dialogue about the prototypes yielded helpful feedback and new ideas.

Both were very enthusiastic about the activity. They believed that the activity made clear what an algorithm actually does, and expected students to be able to relate that to their own YouTube pages. Furthermore, they stressed the importance of the application being visually attractive for students, which was already explicitly mentioned earlier by one teacher. A practical suggestion

was given by one of the experts, who said that Robin's interactions with YouTube could be visualized through icons.

5.3.3.2 Implications

The prototype of the activity was once again met with positive reactions and enthusiasm from the students and experts we showed it to. However, it appeared from these evaluations that the main issue yet to resolve is the lack of difficulty in the activity. The 'storyline' with Milan Knol makes choosing between videos so easy in this iteration that every student could conceivably achieve the maximum score. The concept of recommending similar videos is not extremely hard to grasp - the thought processes that students appeared to be having about what video to recommend already proved that - so the activity being easy is not inherently bad. However, students themselves indicated they would like it to be more challenging, and suggested some ways to achieve that. The next iteration would need to be more difficult.

5.3.4 Fourth iteration

This fourth iteration of the Figma prototype was the last major iteration, and the iteration that received the most amount of feedback: this prototype was shown to and discussed with in total eight students and three teachers.

The primary goal of this final iteration was to increase the difficulty for students and provide some challenge for them. The main 'problem' with the last iteration was the prominence of Milan Knol and the obviousness that his videos were the correct ones to pick. Several relatively drastic changes to the storyline were made to diminish that predictability. First, the amount of information in Robin's profile was reduced to just name, age, and YouTube channels he is subscribed to (Screen 14). This not only more accurately represents knowledge that an algorithm might be fed with, but also requires students to make inferences from Robin's subscriptions to his interests, adding a layer of complexion. Second, we also decided against showing what Robin is currently watching on YouTube (Screen 10), which means that students have to make a decision purely based on Robin's (minimal) profile.

Additionally, alongside Milan Knol, a second interest of Robin was added to the storyline through one of his subscriptions: freestyle footballer Touzani, signifying his interest in football videos. While the first round remained the same (Milan Knol & Minecraft) as the correct option, Touzani now would be the right video to choose in the second round. The third round contained no videos related to Touzani or Milan Knol, but did include a video on football, requiring students to make the inference from Touzani to football. The fourth round forces students to choose between two seemingly correct options, as it contains two videos from the same series starring both Touzani and Knol (Screen 16).

Visually, the suggestion from Mira Media to add icons that show what interactions Robin has with YouTube was directly implemented in the feedback screen (Screen 15). This helps students to understand what Robin's response is at a glance.

Figure 12: Screens 14 to 16.

5.3.4.1 Feedback

This iteration of the prototype was evaluated with two focus groups of students (FG3, five students, and FG4, three students) and in one interview with three teachers (D13, D14 & D15).

Students. Similar to earlier focus groups, we guided the students through the activity, but they had the control over what videos to pick and why. Students from both focus groups liked the activity (*"really fun",* L11) and also believed it to be a nice way to learn something:

> *"You'll have to pay close attention and think about it. You can't just sit on your phone, because then you won't know any of this."* (L15)

The difficulty of the activity had clearly increased compared to earlier iterations. Students were not breezing through the activity anymore, and disagreed with each other on what video to recommend to Robin. Even still, the students remarked that they would like it a bit more difficult. L11 suggested providing even less information about Robin beforehand, so students would have to generalize and figure out what he would like that way:

> "*If you don't know what channel he's subscribed to, what he likes, then you'll have to think: 'he is a boy, he is 13, what do some 13-year-old boys like?' Then it can get harder, you'll have to start using your head more."* (L11)

The intermediary feedback, although generally appreciated, might also have a role in making certain decisions too easy. L11 valued the 'interest arrows' feedback, but said that the texts (as for example seen in Screen 12) might be too telling:

> *"That bit of 'this video of Kalvijn, he is subscribed to these channels', that provides you with more information, and then it gets easier. While you want it to be harder."* (L11)

Even though students said it was not all too difficult, choices for videos were not as easily made as in earlier focus groups. The kind of reasoning we saw in FG2 also emerged in FG3 and FG4:

> "*I think the one from Milan Knol, because he* [Robin] *is subscribed to him. And when you subscribe to a certain channel, you want to see more videos of him."* (L11)

> "*I think it's better, that video with football, because Robin is subscribed to Touzani and Touzani makes videos on football. So maybe he likes football."* (L8)

> *"I think the one from Gamemeneer or Touzani, because Gamemeneer has the same kind of videos as Milan Knol."* (L9)

> "*That one from DylanHaegens, because he* [Robin] *is subscribed to Supergaande. That's more of a comedy, and DylanHaegens also makes Top 10 videos in a funny way."* (L9)

Some students in FG3 however went above and beyond the Milan Knol and Touzani storylines we had devised for this activity. Students started thinking about what videos Robin would find exciting to watch regardless of his personal interests, making choosing between videos a lot harder. In the last round, where both Touzani and Milan Knol appear in videos with a type of challenge, L12 suggested that because Touzani is fitter and more athletic he would probably win the challenge, so the video with Milan Knol would be more exciting to watch. L10 proposed to choose a video not related to Robin's subscriptions, because "*he's there the entire time, maybe he wants to watch something else than just that person"*. L9 also took the thumbnails into account and wanted to recommend the videos with 'clickbait' thumbnails, because *"he* [Robin] *might be impressed by those pictures"*.

While students in FG4 indicated that they would like to see some kind of competition in this activity ("*I think playing against one another is a bit more fun"*, L14), students in FG3 thought the opposite. They expected the element of competition to be rather distracting, "*because you'd just want to win*" (D11). Despite that, they did appreciate the interest score - not to be able to compete against each other, but as a means of finding out if they understood it correctly:

> *"I'd like this, because if I know that I am not that good at it, I can go ask for more instruction from the teacher or a classmate. So I can put more time into it to do better next time."* (L12)

Teachers. The teachers responded positively to the activity, and believed that their students would like to play it as well:

> *"I think the concept is very nice. Those kids will feel like: 'hey, I have power, I have control, I am going to decide what you're going to watch."* (D14)

The activity is not stand-alone - it precedes a class discussion with the teacher - and the teachers also made sure with us that this was the case. In itself, the activity would not be enough to teach the students something:

> "*So then we would need to connect it* [the activity] *with something like: 'and why do they do it? And what is the business model behind it?'. Because then it all comes together."* (D13)

Regarding the difficulty, these teachers shared a different point of view than previous teachers we spoke to. They figured that students getting most choices right is not necessarily a drawback:

> "*I think it makes sense that they get it right. It's fairly easy of course, it's very simplistic, superficial, 'this is what you like and this isn't'."* (D13)

D14 also noted that perceived difficulty is greatly dependent on the class, as *"one class can handle it better than the other",* adding that it is up to the teacher to determine what works best for their specific class.

5.3.4.2 Implications

Even though the students in FG3 and FG4 indicated that they would prefer to be challenged more, their reasoning and behaviour when picking videos to recommend for Robin showed that they are playing with and thinking about the activity as intended. Students were genuinely thinking either like an algorithm, or how an algorithm would decide. It is difficult to increase the difficulty of this activity further without having students making completely random guesses, which we believe would be unwise.

There is however a genuine possibility for a sequel to this activity. That second playthrough could not only include more controversial or political topics, as D9 suggested earlier in the process, it could also conceivably be used to throw students in at the deep end. It could be used to teach (more advanced) students about algorithmic phenomena like the cold-start problem, or differentiate between different kinds of recommendation systems. This however is outside the scope of this project.

Whether or not to include collaboration and competition remains undecided after evaluating the prototype. Students and teachers alike had outspoken opinions on cooperating, collaborating, and competing, but no general consensus was reached. It appears to be primarily based on personal preference and past experience, but further evaluations can be done to find out whether it also impacts learning with this prototype.

One issue brought forward by one project member was that of copyright. It is possible that copyright applies to the videos (thumbnails and titles) that are used in the activity, as they originate from real YouTube-channels. Because one of the strong points of this activity is the usage of relevant and relatable YouTube-channels, videos cannot just be replaced with fictional videos. Moreover, students would not be able to make inferences like they did for videos and channels they are familiar with if videos were fictional.

5.3.5 Summary

From the start and the very first prototype, this activity was received very well by both students and teachers. Both provided valuable feedback that was used to further develop the activity.

The final version was slightly adapted for the final evaluation of Anneleen Janssen; some videos were replaced, but the storyline was generally kept the same. The full version can be found in Addendum 1.

5.4 Activity 3

This section describes Activity 3 (working title: *Bubble Activity*) which came to fruition during a later stage of the project, after Activity 1 was put on hold. Detailed below, aside from the prototype itself and the feedback it received, is the process that generated the idea for this activity. Since the idea was conceived relatively more towards the end of the project than Activity 2, a single iteration could be evaluated with students and teachers.

5.4.1 Origin

When the first activity was shelved (see Section 5.2), only Activity 2 remained. In the meantime, the anti-filter bubble project had produced a directive containing seven distinct themes and lessons to be covered in the anti-filter bubble application. Using these seven lessons as a guiding framework, we conducted a brainstorm session with ourselves to generate new ideas for activities.

The approach we took was similar to the nominal group technique for brainstorming (Dunham, 2006), although slightly adapted to the circumstances, as there were only two of us and the brainstorming session took place online (using Mural). We asked ourselves four questions from the start to guide our thinking:

(1) How can we create an activity that engages students?
(2) How can we create an activity that teaches students something?
(3) How can we create an activity that teachers trust?
(4) How can we make sure the activity is more than a single-use gimmick?

Using these four questions, along with the seven lessons framework as a guideline, we started the first phase of the brainstorming session. We individually generated ideas and wrote them down on individual post-it notes, which were not yet visible for the other. After twenty minutes, all post-it notes were revealed, and we discussed them one by one to clarify the ideas. Finally, we used dot-voting to indicate which ideas we believed were the most promising among them (Figure 13).

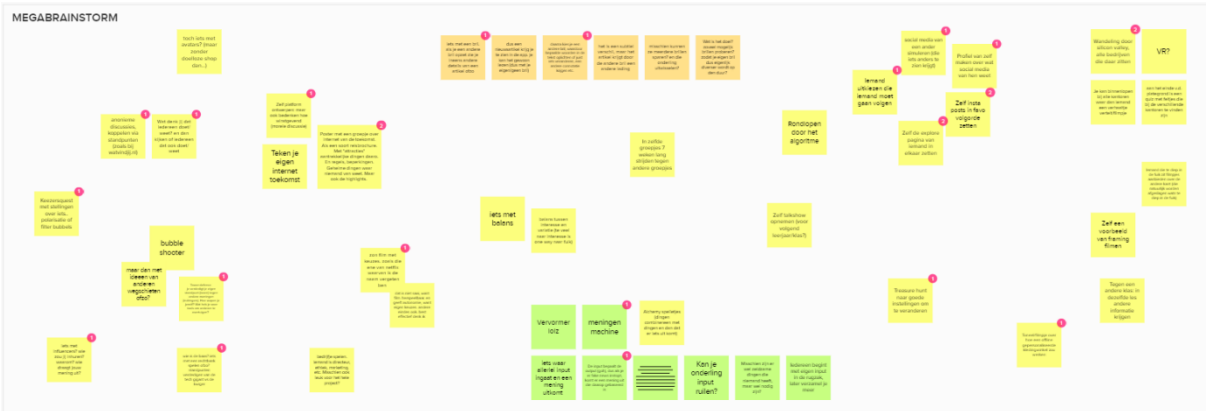

Figure 13: The Mural post-it note wall with the results of the first brainstorm session.

In the next phase, we took six ideas that seemed the most promising, and went on to work out those ideas (see Figure 14). Once again, this was first done individually and with concealed post-it notes. Afterwards, we shared what we had written and discussed the ideas. We also started considering practical obstacles and difficulties here, which we had previously intentionally disregarded.

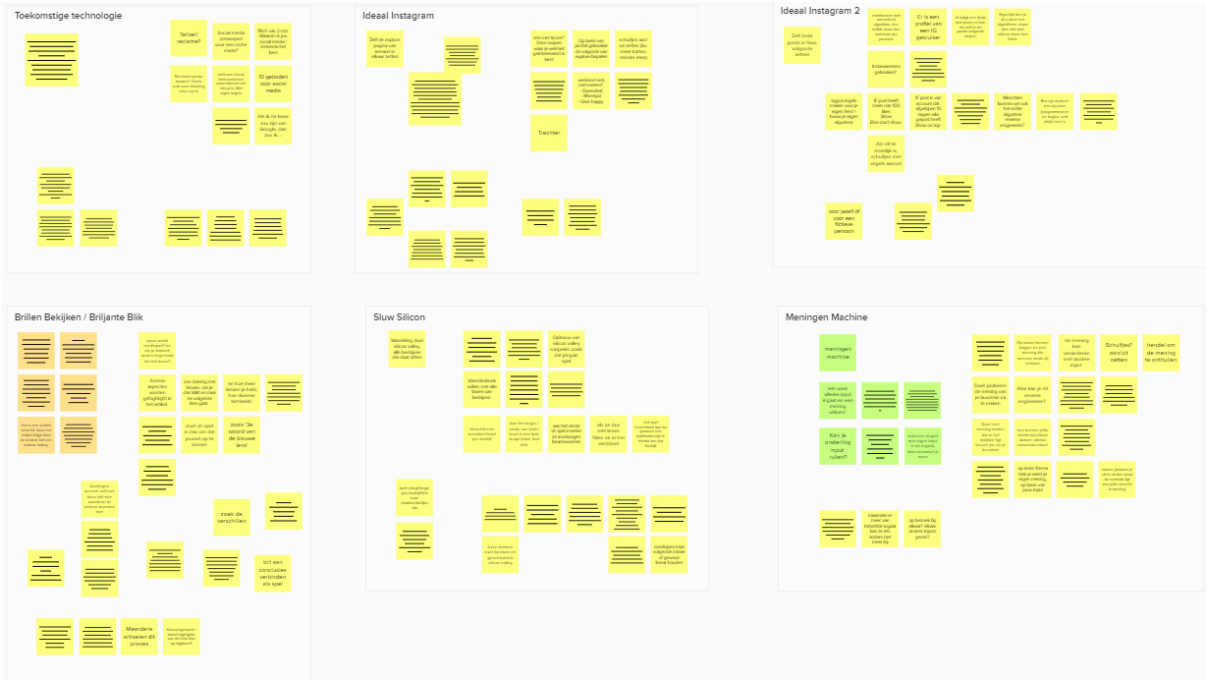

Figure 14: The Mural wall with the six most promising ideas.

Ultimately, we took the idea that appeared to be the best combination of feasibility, engagement, and educational value, and discussed its possibilities, challenges, and key focus areas. This idea was loosely inspired by focus groups that the pedagogy students conducted. To visualize the effect of filter bubbles, they placed videos and other online sources in two paper circles, which represented the two bubbles from two different people. The more the contents of the circles started diverging, the more the bubbles moved away from each other, depicting polarizing effects of filter bubbles.

The general idea of the activity is as follows: in the first part of the activity, players 'walk through' an interactive story using a character that they chose themselves from a fixed set of choices. Every character has their own story, with a fixed end point. During the story, the players make choices as the character and are tricked into believing that these decisions they make during the story decide where their character ends up. However, these decisions have a very minimal impact on the outcome of the story. Every decision that is made ends up in the 'bubble' of the character, which is thus shaped by the choices made by the players: what they decide their character clicks on, what he sees, what he likes or comments on, etcetera.

The second part of the activity commences after players have finished their characters' story. Their character will then enter a conversation with another character, who has experienced an entirely different side of the story. The bubble that the players have shaped during the process

should be used as input for the conversation. Players are led to the insight that their online behaviour, the choices that they make when they are online and their (filter) bubbles can be of influence on their beliefs and view of the world. The intention is that after this activity, a class discussion will take place which is more geared towards understanding the consequences of this. Something important to realize is that the activity should not be teaching students that one certain opinion on a topic is good or bad. They should be learning about the process of forming an opinion, not about the topic at hand. Therefore, the conversation in the second part of the activity and the subsequent class discussion are so important. The activity thus directly caters to Nieuwelink's (2020) three components of knowledge teenagers need in a tolerant, multicultural society. It teaches students that societal issues always involve different perspectives, that tolerating these other perspectives is important, and that sharing your perspective with others is necessary.

Two design decisions had to be made before a first prototype could be developed:

Topic. A topic should be selected around which the story itself revolves. Characters will end up at a different end point, which means that the public opinion on the topic must be somewhat contested - if all characters agree with each other because there are no multiple sides to the story, the activity has no value. However, opting for a topic that is too controversial possibly undermines the safety of students in class and their freedom to speak their mind (such as racism or gender identity). Several topics were considered, such as climate change, vegetarianism, and violence in games. Ultimately, the topic chosen to be in the activity is the presence of (young) children in family vlogs and videos. This subject was brought forward by students themselves during focus groups, indicating their involvement with the subject. It is a topic that many students will recognize, but most will not have a very strong opinion either way, which makes it the ideal matter for the storylines in this activity.

Characters. The students will play through the story with a character that they chose themselves from a fixed set of possibilities. We decided to not let students create their own personalized character and play as themselves in order for them to be able to dissociate from their own opinions and ideas about the topic. However, we did choose to make the characters as relatable as possible for them - same age and same habits in social media.

Ultimately, after generating a number of sketch drawings to visualize the idea, we made the decision to start working on the prototype and show this to members of the anti-filter bubble project team for initial feedback.

5.4.2 Prototype

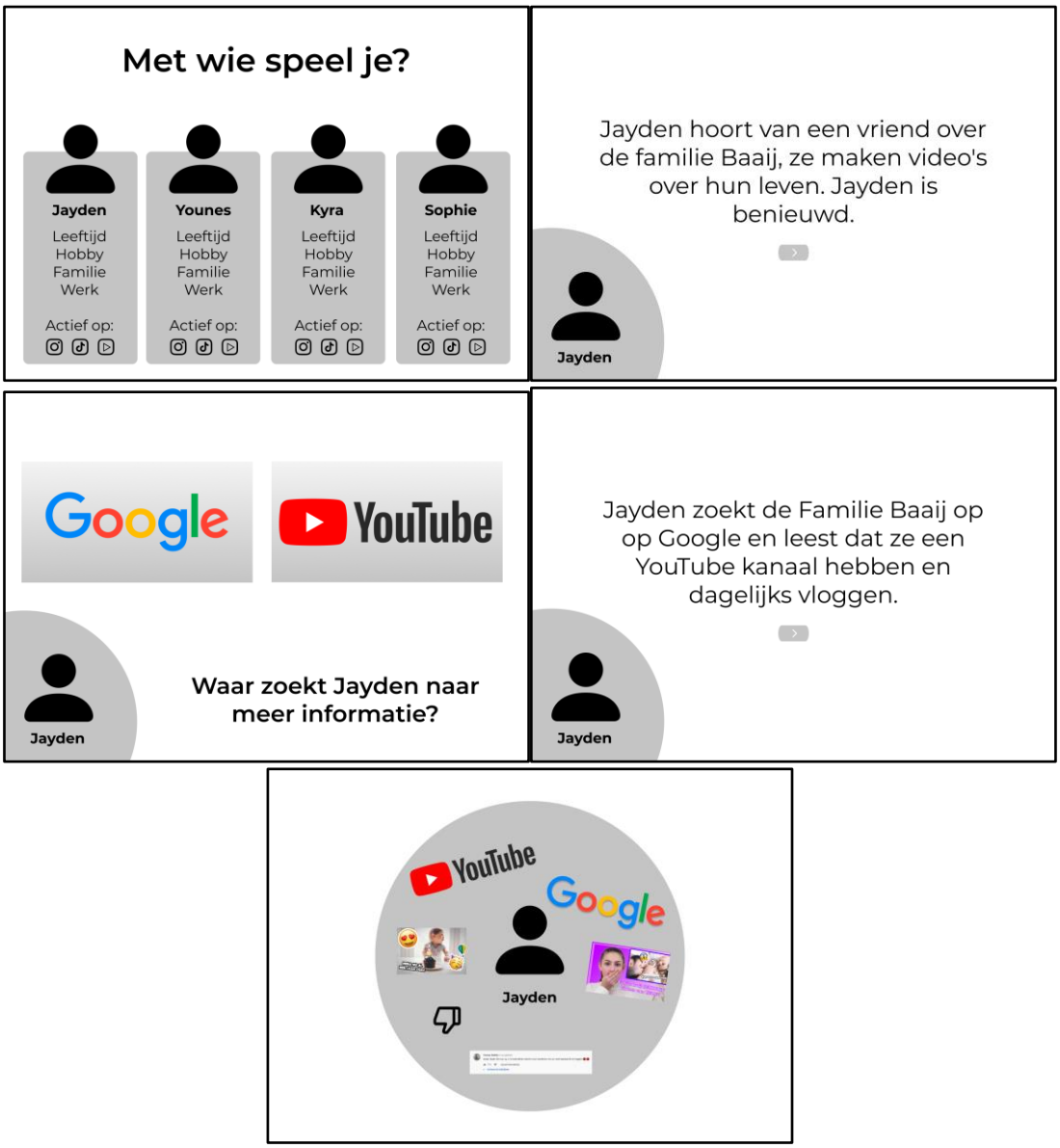

Figure 15: Screens 1 to 5 of Activity 3 (left to right, top to bottom).

In the prototype, we developed two of four characters' stories (Screen 1) to be able to show the concept of the activity. The two we developed first are those of Jayden and Sophie. Jayden's story is that he first hears of the (fictional) family Baaij from a friend (Screen 2). He then goes to search for more information, either through Google or YouTube, which is the first choice for the players (Screen 3). In Screen 4, the player has chosen for Jayden to use Google. The full storyline flowcharts are displayed in Figure 16 and described in text in Appendix I. Eventually, Jayden discovers a number of comments that say that vlogging is bad for children, and after looking this up online, Jayden is more and more convinced of the harmfulness of vlogs for children. The storyline ends with Jayden himself posting hateful comments against the Baaij family. Screen 5 displays a possible bubble for Jayden at the end of the story. In the meantime, players that have played with Sophie will have experienced something else. Sophie has become a fan of the Baaij

family and believes that people like Jayden should mind their own business. Sophie's storyline is also visible in Figure 16 and written out in Appendix I.

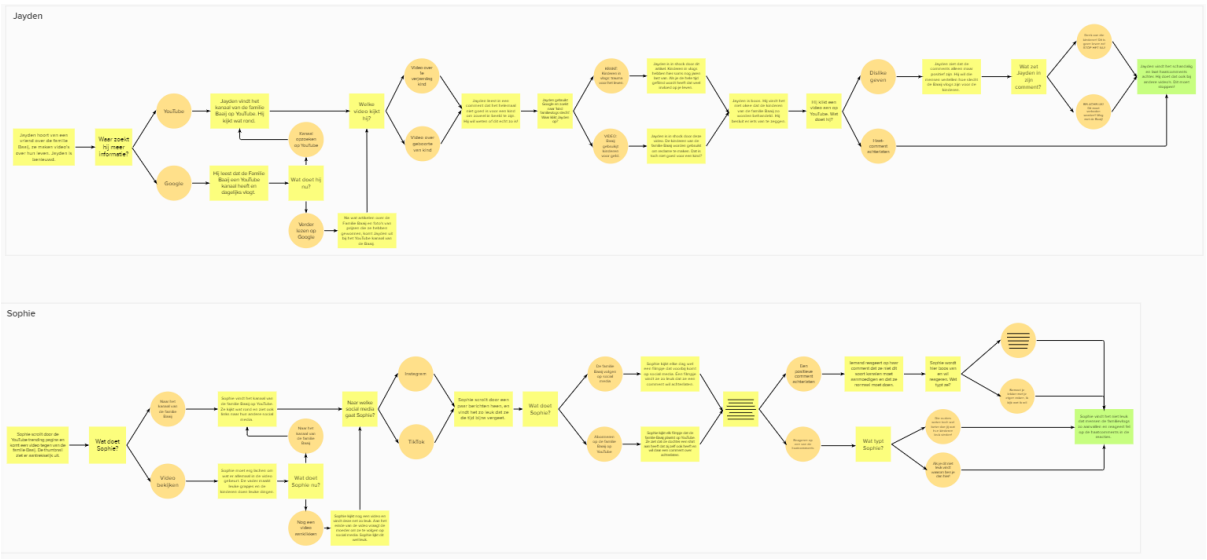

Figure 16: Jayden's (top) and Sophie's (bottom) storylines. Yellow shapes are texts and choice prompts displayed to players, the orange circles are the two options per choice, and the green rectangle is the final end point.

We also worked out the second part of the activity in the prototype. Students that have picked and played with Jayden will be chatting with students who have played with Sophie. We opted against providing players open text fields, but will instead offer players two possible options to choose from (see Screens 6 and 7). This guarantees that the conversation stays within limits and on topic. At fixed moments in time, players can drag an item from their bubble to the chat, to explain why they believe certain things and to support their opinion. Of course, the opposite player can do the exact same to show where their character got their beliefs from.

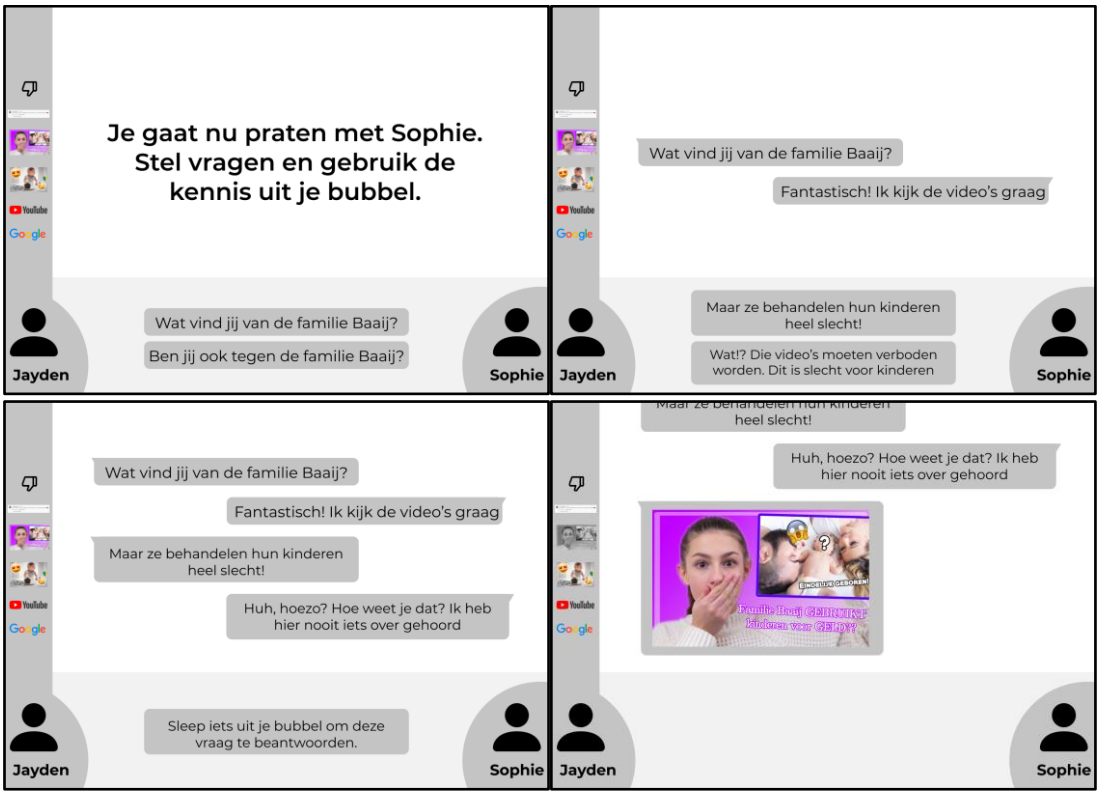

Figure 17: Screens 6 to 9 (left to right, top to bottom).

The first initial feedback from project members was positive and enthusiastic. They said the activity comprehensively shows the underlying 'mechanisms' of filter bubbles. However, as mentioned above, it should be clear for both students and teachers that the process is the learning experience, and not the outcome - students are learning about filter bubbles, and not about family vlogging. Additionally, they emphasized once again the importance of the class discussion that follows the activity. There, students should realize the how, why and real-world consequences.

5.4.3 Feedback

Given that the idea and subsequent prototype were conceived relatively late in the project, there were limited opportunities for evaluation. Eventually, we were able to show the prototype to two focus groups of five students in total (FG5, FG6) and three teachers (D13, D14, D15).

Students. Students responded positively to the prospect of their choices having an impact on the story and aroused their curiosity:

> *"Wait, can I play again and make different choices? I want to see what happens. What happens in other endings? I want to know. Is there another ending?" (L16)*

Some drew a comparison with Netflix's interactive movies such as *Bandersnatch* or *Minecraft Story Mode,* which they said was very enjoyable. However, students indicated they would have

liked to see a longer story with more opportunities for choice. Two options per choice was enough, but they wanted "*more questions, and more endings, so that it will spread, branch out*" (L19).

Regarding the story itself, FG5 did not perceive it as realistic, especially the opening. One student suggested embellishing the story a bit more:

> *"There just needs to be a good trough line in the story, like 'he comes home from school, gets some food, goes on his phone, and suddenly sees a …'"* (L16)

The topic of children in family vlogs was met with mixed feelings. Some students did not like it as much, while others were interested in the topic and already had some opinions on the issue themselves. Suggestions for other topics included public opinion on influencers, and one student proposed including a picker wheel for topics at the start, which would decide what topic the story would be about (because *"what child doesn't like to gamble?"*, D16).

Additionally, one student in FG5 also suggested that they should be able to personalize their characters, which was met with approval from the other students in that focus group. Their suggested personalization was relatively superficial, such as customizing their (second) name to distinguish them from other Jaydens or Sophies in class, or adding different clothes such as hats or shoes.

Students disagreed with each other when asked how long this activity could or should be played for. One said somewhere between 15 to 20 minutes, another said 5 minutes would be enough. L19 took another approach and estimated the duration from the perspective of a teacher:

> *"First, we'd wait until everyone has settled, and some children would not understand it completely. It would take a little bit longer. I'd give it fifteen minutes. Ten minutes for explaining, and five minutes to work on it."* (L19)

In terms of difficulty, there is no 'good or wrong' aspect to this activity like there is in Activity 2, but students also did not have any trouble with making choices or with reading the texts in between ("*because you don't have to think about it that much, and you can just choose"*, L20).

The focus group ended with us showing students the second part of the activity, the chat between students that played with different characters. It was immediately met with a lot of audible enthusiasm from L17 ("*because I've never seen anything like this"*), and other students also watched with interest. L18 drew the comparison with *Yarn*, a mobile application in which players chat with fictional characters. L19 liked the possibility of dragging items from their own bubble into the conversation, to provide evidence for their opinions.

We discussed two design issues for this chat that needed resolving with the students. One of them was the decision we made to provide students with two predetermined options instead of an open text field. Students generally agreed with this, mainly because they had the ability to foresee what would happen if we provided them with open text fields:

> *"A question for you: what do you think would happen if all words could be used?"* (L19)

"*That doesn't make any sense, because then we'd be able to just go off topic."* (L16)

The second design issue is the distinction between chatting with another classmate (who played the activity with another character), or chatting with the computer (who would give answers as if it were the other character). Students primarily indicated they would prefer talking to a classmate (*"because with a computer, I think, it's just fake, it feels fake"*, L20). Interestingly, they were also able to take on the view of a developer, and mentioned chatting with a computer - in the case of open text fields - would be near impossible to create:

> *"That's impossible, because the system - if you type in another word that isn't related at all to any of this, the system will explode"* (L17)

L19 mentioned that their preference of talking to a computer or a classmate would depend heavily on who that classmate would be:

> *"With her* [L20], *if it'd be her, it would be interesting. But with them* [pointing to the rest of his class], *you don't know if they know what they're dealing with."* (L19)

Students also gave feedback on the aesthetics of the prototype, and recommended the use of both colours and sound, as *"otherwise, it would be a bit boring"* (L16).

Teachers. The teachers' first response was positive: they especially liked the interactive nature of the activity. The topic of children in family vlogs was also appreciated, and teachers immediately started thinking about ways they could apply it to real-world situations. Teachers saw possibilities for the subsequent class discussion that this activity allows for ("*I think that it* [the activity] *does provide input for a good conversation",* D14). According to them, the activity itself would already get students to think about their online behaviour, but the following class discussion would make it more 'real':

> *"These children, on average* [...] *will realize very quickly: this game is not quite real. They're trying to teach me something here. But if they then see 'this is really how it goes in the real world', then it becomes real for them."* (D14)

The teachers also saw some difficulties regarding the story. Echoing the opinion of L16 above, D13 was worried that students might not consider the storyline to be very realistic. Another issue was raised by D14, who worried that maintaining actuality in this activity over the years or even months would be difficult.

While the teachers, unlike the students, did not suggest adding more depth or length to the story, they did propose to let students play through the story multiple times. That would make the effect of online behaviour on forming an opinion get across better, according to them.

Worries about the amount of text throughout the activity were dismissed by the teachers:

> *"I don't think the amount of text is too bad. It's always on one slide, and that helps a lot. If they see a really big chunk of text, they will disengage, but if it's just a few sentences, they'll read that."* (D13)

D15 suggested adding emphasis to the important keywords in texts, so that even the students who are not as engaged (*"some will click through it faster"*, D13) read the important keywords.

5.4.4 Implications

Even though there were relatively limited opportunities to gather feedback on this activity, the five students and three teachers we spoke to provided useful and insightful comments.

The activity provoked positive and enthusiastic reactions. Choosing your own story clearly connected well with the students, who were eager to try out different storylines and wanted more choices and a more developed storyline. Teachers also saw ample opportunities for the following class discussion and believed that the activity really could lead students to rethink their online behaviour.

The reaction to the topic of children in family vlogs was as expected. Some students liked it, some would have rather had another topic to play with. There is not one single topic that appeals to every student, but it should be meaningful and connect to the real world, which it did, as both students and teachers were already connecting the storyline to real world examples during the focus groups and interviews. One student suggested to add a picker wheel that would decide on the topic before the activity. While a gambling element in an educational application might not be appreciated by most educational institutions, it is a viable possibility to develop multiple storylines on different topics and supply them with the activity. A teacher, or even the class itself, could then decide on what to play with.

Conclusively, this third activity was well-received and provides plenty of opportunities for further development. The full final version can be found in Addendum 2.

5.5 Overall application structure

As mentioned in the description of the brainstorm session for the third activity, the anti-filter bubble project produced a directive with seven distinct themes and lessons that ought to be covered in the application. While this setup helped to organize and structure the up until now relatively vague concept of the anti-filter bubble application, the seven lessons were still loosely connected to each other and an overall sense of 'why' was still missing. A general theme or 'story' that runs throughout the entire application should help to connect separate lessons to one another.

Based on one idea that was produced during the brainstorm session as described in Section 5.4.1, we devised an overarching structure for the application. The idea in question, originally imagined for one lesson only, was to let students 'design' their own social media timeline with certain goals in mind. These goals could for example be to make as much money as possible, to increase diversity of posts, or to make users as happy as possible with their timelines. We realized this could just as well work as a general theme throughout the entire application.

Students would get to make their own social media platform during these seven lessons. They would work towards that 'big end goal', as teachers recommended us in the first pre-study. After each lesson, they would have to make choices for their platform based on what they had just learned in that lesson, e.g. decide what recommendation algorithm runs on their platform after doing Activity 2 (Section 5.3). This helps to provide students a reason for learning about filter bubbles, algorithms and social media: they need to make well-educated choices for their own platform. Their choices would impact three metrics: user base size, monetary profit, and diversity. Connecting these three 'scores' to students' choices would introduce an element of competition and a further reason for students to take the matter seriously. Finally, creating your own platform also offers possibilities for creative personalisation (such as logos, colours, or lay-outs), which we noticed in focus groups appeals to students, and also further gives them a sense of ownership.

The next section describes the first prototype that was developed for this overall structure idea, and the feedback that we received on it.

5.5.1 First iteration

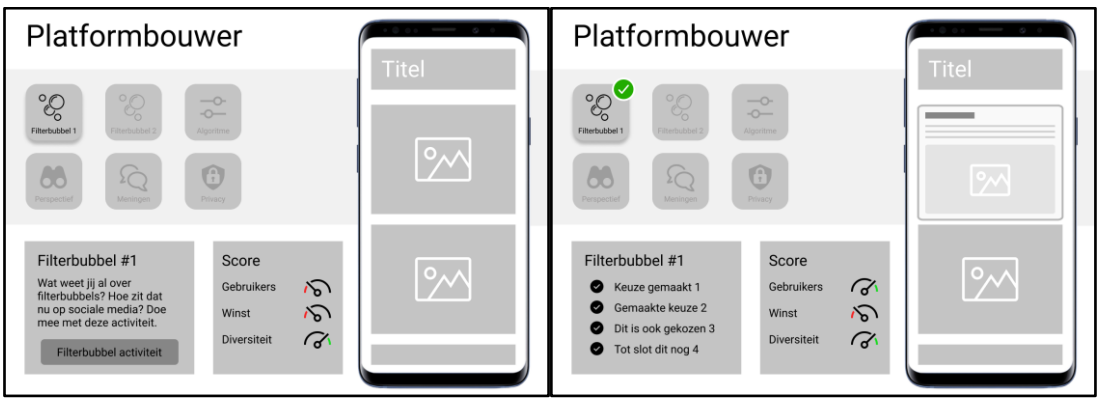

Figure 18: Screens 1 and 2 of the overall structure.

The ‘platform builder’ would be the interface from where students start and end all lessons. All lessons would have a similar setup: (1) start with an intro from the teacher, (2) do the activity, (3) class discussion and (4) make choices for the social media platforms.

Screen 1 and Screen 2 (Figure 18) display the platform builder before and after students have gone through an activity. Visible are the six different lessons (the seventh was envisioned as a final concluding lesson and therefore not included), with all lessons except the current one greyed out, to indicate to students where they are in the process. On the right side, a smartphone is visible that displays the social media platform that they are developing. Before going through an activity and making choices for their platform, elements are still greyed out or not visible, but after students have made choices, certain parts of the platform are filled in. In the bottom-middle, the three gauges for the three metrics are visible. For illustration purposes, the user base size gauge changes from low to high between Screen 1 and Screen 2. These gauges were inspired by the Bad News Game[14] wherein players develop a fake news platform and their performance is also monitored by a gauge that measures the number of followers.

The bottom-left of the prototype contains specific information on the lesson that the students are about to begin with (Screen 1) or on the choices that they have just made after the activity (Screen 2).

5.5.1.1 Feedback

This version of the overarching structure idea was primarily evaluated with experts within the anti-filter bubble team and Mira Media, but was also shown to D12 and D13.

Pedagogy. The first response from the pedagogy team we showed the prototype to was positive. They thought the concept connected well to the needs of vmbo students and appreciated this approach for connecting the seven lessons with each other. They also unearthed several crucial questions and issues however. The metrics and the scores that would be connected to them were the main point of concern, especially that of diversity. According to them, diversity could not be captured in one single metric, as there are many different sides to diversity. They feared that the platform builder would make some concepts too superficial and that the simplistic versions of some concepts such as diversity would be incorrect. Another worry was the moral judgement that would come with certain choices, as the metrics would add objective ‘good or wrong’ value to the platforms of students.

It was also suggested to provide students with the opportunity to return to previous choices, and change them if they wanted to. Reason being that if students learned new information in later lessons, they may not agree with their earlier choices, and continuing with something you know is wrong would not be motivating for them. Additionally, in accordance with what we had in mind, they also proposed to turn the last (seventh) lesson into a session where students could present their own platform and the choices they made for it.

[14] https://www.getbadnews.com/

Mira Media. The two experts at Mira Media also approved of the idea and its prototype. They confirmed that working towards one end goal, in this case their own platform, would help to motivate students.

Entirely opposite to the project team members, the two experts appreciated the diversity metric, but were more worried about the other two (user base size and monetary profit) metrics. The diversity metric would, according to them, provide a great start to a class discussion about what high or low diversity on a platform could mean. However, how students would value user base size or monetary profit differs per school and student, making it a highly subjective metric; students could for example believe that a small user base is better, because the platform would be more exclusive to them. The experts suggested adding more objective diversity-related metrics instead, such as inclusivity, but also realized that the metrics should be straightforward and easily explainable to (vmbo) students.

Teachers. The two teachers that were shown the prototype (D12 and D13) responded favourably and enthusiastically towards it. They appreciated the non-traditional approach of teaching:

> *"Didactically, it fits our school really well, in the sense that you're not going to act all righteous and boringly explain why a filter bubble is wrong. They're really working on it themselves."* (D13)

> "*This app, in my opinion, in which all the pieces of that big story are revealed one by one, is actually already very helpful and supportive, in order to talk about the big story at the end."* (D12)

D13 also mentioned that because students need to make choices for their own platform, they will have to think about the topics at hand, creating the necessity for learning that was not apparent throughout the application before.

The metrics also caught the teachers' attention, as they saw plenty of openings to a class discussion, especially for the diversity metric:

> *"That diversity parameter, we as teachers feel 'that really should be in the green, because that's important'. But that's another context that needs to be addressed. Because why is it important to have that diversity? That will still be a bit abstract to them."* (D12)

Finally, we also asked the two teachers if these platforms should be built individually, or that students should work in groups. Practically, they said groups would be more convenient, as judging or discussing 25-30 different platforms and differently made choices would be a lot harder than 'just' several groups. On the other hand, they also recognized that this would somewhat remove the feeling of ownership over the platform, while "*the design really invites you, like 'I want to work on my own app'"* (D13).

5.5.1.2 Implications

A lot of attention during the 'evaluations' went to the metrics, what they signified, what value the students would attach to them, and if they were or were not oversimplified. It appeared that they

were a feature with a number of (hidden) catches, and deserving of an extensive review in itself. Because this falls outside the scope for this thesis project, we decided together with the project team to abandon the metrics for now, and make further developments to the prototype without that specific aspect to it.

There also is no clarity yet on the issue of collaboration or cooperation while building the platform. However, as this is something that would not require much adaptation to the prototype or final application, there is no urgent need for any design decisions on whether to include cooperation.

The next step is to embellish the prototype, and add screens that display how the process of making choices for the platform would go.

5.5.2 Second iteration

There were several major changes to this second iteration of the prototype. Primarily, the overview screen (Figure 18, Screen 1) received a visual overhaul (Figure 19, Screen 3), with more focus being placed on the ongoing lesson instead of on the entire lesson series. This makes it more clear for the students what the next step will be in class, and removes the unnecessary information on future lessons.

Second, the process of making choices for the platform was added to the prototype. Screen 4 shows the four choices that students would be able to pick from in this sample question. When hovering over one of the options, the squares reveal more detailed information on that choice (Screen 5). This text is relatively superficial - only after students have made their final choice, the consequences for their platform are revealed (Screen 6). Then, in order to make those consequences more insightful for students, we also added two overviews of their social media platform. The foreground overview (Screen 7) shows direct consequences, in this case certain comments being removed and users being banned because of making derogatory comments. The background overview (Screen 8) displays the more 'algorithmic' choices that students have made, such as what their recommender system does or in this case, if the system filters out offensive comments.

Conceptually, and primarily in consultation with other project members, a 'storyline' was added to the framework of the application, providing more structure and meaning to the idea of building a social media platform. This storyline consists of three phases: in phase 1, students are users of a platform and learn about basic concepts surrounding the filter bubble. In phase 2, students take on the role of a (low-level) developer and make the first decisions for their platform, mostly of algorithmic nature. Phase 3 concludes the lesson series, and has students making more ethical or higher-level choices for their platform.

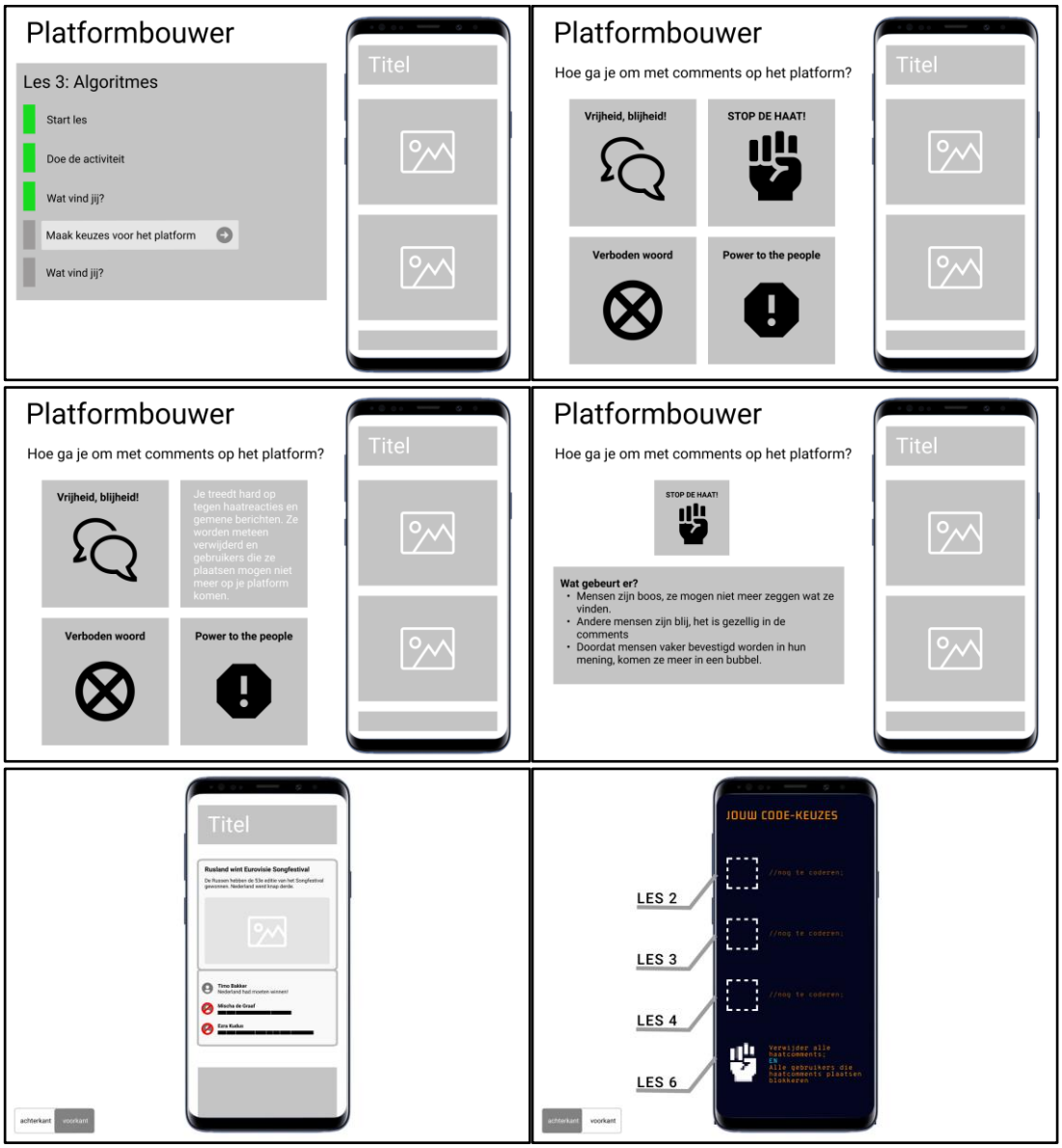

Figure 19: Screens 3 to 8 (left to right, top to bottom).

5.5.2.1 Feedback

This iteration of the prototype was shown to three teachers. One of them, D13, had also seen the previous iteration, giving them the opportunity to give specific feedback on the improvements. Because of this, the conversation mainly revolved around the process of making choices for the platform.

Both D13 and D14 expected the students to be able to enter discussions about the choices they would be making for their platforms, or at least about the underlying principles. Teachers saw a lot of value in being able to connect to those bits of theoretic principles through these choices:

> *"When you discuss fundamental rights, or freedom of speech versus respecting each other, well, they'll have different opinions about that. So I think that's super interesting, I like it."* (D13)

D13 also recommended taking a further step towards students' real lives, by for example comparing their choices with how things are currently implemented in Instagram or TikTok. D14 suggested providing teachers with a 'dashboard' that would give insight into what choices their students are making or have made.

Regarding the prospect of collaboration, D14 mentioned one big drawback they foresaw if platforms were to be built in groups:

> *"The big disadvantage of groups* [...] *is that there's one student behind the controls. What will the others be doing?* (D14)

5.5.2.2 Implications

Although the feedback on this second prototype was relatively limited because of the priority that Activity 2 and Activity 3 had, it confirmed for us that the process of making choices was implemented correctly. Teachers even saw possibilities for connecting class discussions to the reasonings of students for making certain choices.

The suggestion of D14 to include a dashboard for teachers closely relates to Requirement 2, which states that teachers should have a clear overview of what students are doing within the application. While it is not yet part of this final iteration of the prototype, we believe that granting the teachers access to such a dashboard would be helpful in providing them that overview.

5.5.3 Conclusion

We consciously decided not to evaluate the prototype of the structure with students. Two reasons contributed to this decision: first, given the relative scarcity of available focus groups and students, we opted to prioritize Activity 2 and Activity 3, as these were and are the primary focus within the project. Second, contrary to both those activities, this structure does not immediately provide students with a 'playful' experience - students would have to react to us showing them a couple of still images instead of them being able to play through it themselves. It is important that the idea is eventually properly evaluated with students as well, but such an evaluation is better suited for the context of a real lesson, ideally preceded by an activity and a class discussion.

The development of this overall structure idea was slightly beyond the scope of this thesis. Work on the structure was therefore halted here, as the development of the storyline and further implementation of the choices was transferred to other members of the anti-filter bubble project team. Nevertheless, this overarching structure has not only proved to create a solid connection between different parts of the anti-filter bubble application, but it also provides students with a more tangible reason for learning.

5.6 Summary

This chapter has described the design phase of this thesis, in which multiple prototypes for multiple activities and ideas were developed, evaluated, and improved in multiple iterative cycles.

Activity 1, in which students had to compare their own recommended videos on YouTube with classmates, was abandoned relatively early in the process. Initial feedback from teachers and students was not unfavourable, but also not necessarily encouraging. While teachers appreciated some of the key elements, such as the direct comparison with classmates, they also foresaw a number of obstacles should it be used in class, such as the possibility of privacy violation. Along with the negligible role for the application in this activity, this was the main reason for placing further development of this activity on hold.

The second activity was received with considerably more enthusiasm among both teachers and students. By letting students take on the role of the YouTube algorithm, they are granted insight into the inner workings and reasonings of the recommendation system. The iterations of the prototype saw improvements to its difficulty, storyline, and relevance for students, among others.

The third activity was conceived after an extensive brainstorming session, which eventually also produced the idea for the overarching structure of the application. The third activity has students go through an interactive story with a character of their choice, after which they will enter a conversation with someone who has played with another character. The activity introduces students to the influence of online behaviour on forming an opinion and filter bubbles. While this activity did not undergo as many iterations as Activity 2 and has some flaws and design decisions yet to be resolved, it shows a lot of potential and was positively received by both students and teachers.

To connect all (seven) lessons with each other, we also developed prototypes for an overarching structure of the anti-filter bubble application. Students will build their own social media platform through making choices for their platform using the knowledge they have gained during the activities and subsequent class discussions. The teachers that were shown the structure responded positively, and it is now adopted in the anti-filter bubble project and being further developed by other project members.

Conclusively, we have developed three prototypes for activities and demonstrated that these activities have great potential to be used in the anti-filter bubble application. The next step in this thesis is the final evaluation on one aspect of the design that emerged during the focus groups. Chapter 6 will discuss this evaluation and its results.

6. EVALUATION

In the filter bubble activity (Activity 3, Section 5.4), players gradually create a bubble. In the earlier prototypes, this was implemented within the activity as a digital representation, accessible through the application itself. However, in focus groups conducted earlier, also by the pedagogy team, a similar activity was performed on paper. The question arose if either a digital or physical implementation of the bubble would be more engaging or effective.

As most educational materials are either fully digital (such as online teaching methods) or 'tangible' (such as traditional books and exercises), a physical addition to a digital game would be relatively unique. However, Terrenghi, Kirk, Sellen & Izadi (2007) suggested that such a hybrid physical-digital UI may be more suitable. They posit that the "simple mimicking of physical space through graphical representation", as in for example using a mouse to drag choices made to the bubble, "may not be sufficient to encourage interaction [...] like the physical world" (p. 1164). Other academic research has also made a case for tangible interfaces, as they have been shown to improve efficiency and effectiveness of learning (Antle, Droumeva, & Ha, 2009; Zuckerman & Gal-Oz, 2013) and make collaboration with others easier (Miglino, Di Fernando, Di Fuccio, Rega, & Ricci, 2014; Zuckerman & Gal-Oz, 2013). It has been hypothesized that this is because tangible interaction is considered to be more inviting, and because people are inherently accustomed to physical interactions with the environment, making it more intuitive and 'easier' (Piper & Hollan, 2009; Terrenghi et al., 2007).

However, the filter bubble activity would not be fully tangible - the bubble would be merely a tangible element of a digital experience. Having to switch between tangible and digital interaction could disrupt the state of flow of students, while that flow is important to maintain engagement and motivation in the game-based learning environment (Admiraal, Huizenga, Akkerman, & Ten Dam, 2011; Hsieh, Lin, & Hou, 2013). Inversely, the activity might have the students in such a state of flow that they will not even attend to the physical artefacts: earlier research shows that in a learning setting, it is difficult for students to redirect their attention from a mobile device (Eliasson, Pargman, Nouri, Spikol, & Ramberg, 2011). Furthermore, interacting with and moving digital objects, like the choices in the bubble, "is an attractive human-computer interface that definitely 'captures' the children's [sic] attention" (Miglino et al., 2014, p. 90). There are also the practicalities to consider, as having to distribute and possibly collect paper artefacts for each and every student is time-consuming and distracting for students (Liao et al., 2007).

The purpose of this study is therefore to evaluate two possible implementations of the filter bubble activity: either the 'filter bubble' that keeps track of the players' choices is located within the application, or it appears as a physical artefact (on paper). The following questions will be answered with this evaluation:

> Q1: Does the presence of a physical artefact improve the experience of the filter bubble activity?
>
> Q2: Does the presence of a physical artefact improve knowledge and awareness of filter bubbles?

Section 6.1 contains the methodology as intended for the evaluation. However, an unforeseen lack of participants caused an uneven distribution between conditions. As a result, I decided to consult two teachers as experts to corroborate findings and impressions from the evaluation below. The methodology for these interviews is presented in Section 6.2.

6.1 Method for students

This study had a between-subjects design and consisted of multiple evaluation sessions at secondary schools. One group of subjects played the activity with the physical bubble (condition A), the other had access to the digital bubble in the application (condition B). To test whether possible improvements in knowledge on filter bubbles is different in each condition, a filter bubble knowledge test was administered both before and after doing the activity. Additionally, to determine whether the experience of the activity is dependent on the presence of a physical artefact, a shortened version of the GEQ (*Game Experience Questionnaire:* Poels, De Kort & IJsselstijn, 2007) was completed by the participants after playing the activity. Finally, a short form of the DFS-2 (*Dispositional Flow Scale*: Jackson, Martin, & Eklund, 2008*)* was used to determine if there is a difference in flow between the two conditions. Further information on these scales can be found in Section 6.1.3 of this methodology. The scores from these questionnaires, along with the differences between pre- and post-study test scores, were the dependent variables of the study.

The study was performed together with Anneleen Janssen, but the sessions were split in half to accommodate for both our theses' evaluations. The sessions were also recorded (audio only), and during the activity, participants' (inter)actions in and with the application were observed and documented. After playing the activity, a discussion on their experience was facilitated by presenting participants with some talking points:

(a) What do Jayden and Sophie think about children in family videos? How is it possible that Jayden and Sophie have different bubbles? (*on filter bubble knowledge)*
(b) How did you feel about managing your own bubble? (*on experience)*

Four students (3 male, 1 female; also see Table 5) participated in a pilot study. This pilot helped to solve a number of issues. In this pilot study, questionnaires and knowledge tests were distributed on paper; I decided to digitize these for future evaluation sessions because of practical benefits. Additionally, the post-activity knowledge tests were distributed here right after the discussion; students were mostly confused why they had to fill out the same test mere minutes after doing the pre-activity knowledge test. The procedure was therefore slightly adapted so that the pre- and post-activity knowledge tests were at the very beginning and end of the entire evaluation study, instead of just the activity.

6.1.1 Participants

Participants were selected from schools that were participating in the anti-filter bubble project. They were between 12 and 14 years old and in first or second grade of vmbo. This study was designed to have at least 10 participants in both conditions, implying a minimum of 20

participants in total. Ultimately, 10 students participated in the study (3 male, 7 female; 7 for condition B, 3 for condition A). One session (G3) took place online through Microsoft Teams, as their school did not yet allow external visitors because of their Covid-19 measures, which was also the direct reason for the skewed participant distribution among the conditions. One student (P14) was scheduled to participate as well and was assigned a participant ID, but did not show. Full details can be found in Table 5.

Some of the students already participated in earlier focus groups, but none of them had seen Activity 3 before. For the sake of transparency, those students' ID for the design phase are also included in Table 5.

Table 5: Participating students in the evaluation study. 'Kopklas' is in between the last year of primary school and the first of secondary school.

Group	Condition	ID	Design phase ID	Gender	Grade	Located in
Pilot	*B*	*P1*	-	*M*	*1*	*Utrecht*
		P2	-	*M*	*1*	*Utrecht*
		P3	-	*M*	*1*	*Utrecht*
		P4	-	*F*	*1*	*Utrecht*
G1	B	P5	L11	F	Kopklas	Amsterdam
		P6	L12	F	Kopklas	Amsterdam
		P7	L8	M	Kopklas	Amsterdam
		P8	L10	M	Kopklas	Amsterdam
		P9	L9	M	Kopklas	Amsterdam
G2	A	P10	-	F	1	Utrecht
		P11	-	F	1	Utrecht
		P12	-	F	1	Utrecht
G3	B	P13	L13	F	1	Utrecht
		P14	*L14*	*M*	*1*	*Utrecht*
		P15	L15	F	1	Utrecht

Informed consent was collected from students and their parents beforehand. Participants were pseudonymized and both the audio recordings and transcripts were stored in the University's secure data deposit Yoda.

Participants were not compensated with a monetary reward. Due to COVID-19 measures in place, I was also unable to offer any refreshments to participants.

6.1.2 Procedure & materials

Before the start of the session, participants verbally re-confirmed their permission for recording the conversation. Their demographic details (age and gender) were gathered on the first page of the questionnaire.

The chronologic procedure of the study, alongside the materials needed, is outlined in Table 6 below.

Table 6: Chronologic procedure of the evaluation study including the materials, apparatus and time scheduling.

Phase	**What is happening**	**Materials**	**Time (min)**
Preparations	Participants will be informed about what they will be doing. Collection of demographic data.	Recording device	1
Pre-evaluations	Participants individually fill in the test on filter bubble knowledge.	Filter bubble knowledge test	3
Evaluation Activity 2 (optional)	*Anneleen Janssens evaluation of Activity 2 takes place either before or after my evaluation.*	-	-
Explanation	Participants are introduced to Activity 3, and are explained what to do with the bubble.	Prototype Laptop / Computer Paper bubble + items	1
Activity	Participants individually perform the filter bubble activity, either with the physical or digital bubble.	Prototype Laptop / Computer Paper bubble + items	4
Discussion	Based around pre-defined talking points, participants talk with me about filter bubbles and their experience of the activity.	-	4
Post-activity	Participants fill in the GEQ and the DFS-2.	GEQ, DFS-2	5
Evaluation Activity 2 (optional)	*Anneleen Janssens evaluation of Activity 2 takes place either before or after my evaluation*	-	-
Post-evaluations	Participants fill out the filter bubble knowledge test.	Filter bubble knowledge test	3
Closing	If there is sufficient time left, participants can ask questions or provide comments on the prototypes.	None	-

6.1.3 Analysis

This section describes the quantitative and qualitative analyses that were performed on the data that the sessions with students produced, along with further information on the GEQ and DFS-2 that were used.

6.1.3.1 Experience

The short form of the GEQ consists of seven different constructs, each corresponding to two items in the questionnaire. They are:

- *Competence* (items 2 + 9),
- *Sensory and Imaginative Immersion* (items 1 + 4),
- *Flow* (items 5 + 10),
- *Tension* (items 6 + 8),
- *Challenge* (12 + 13),
- *Negative affect* (items 3 + 7) and
- *Positive affect* (items 11 + 14).

Construct scores can be derived by averaging the two item values, which can range from 1 to 5. The full 14-item questionnaire translated into Dutch can be found in Appendix J.1.

Not enough participants took part to perform parametric tests, thus the non-parametric Mann-Whitney U test was performed to determine the significance of any difference in construct scores found between the two conditions[15]. For all statistical tests, a significance level α of 0.05 was assumed. Additionally, the construct scores that arise can be examined separately and conclusions about the students' experiences can be drawn from the scores, especially combined with qualitative observations and utterances (Section 6.1.3.4).

6.1.3.2 Flow

The short form of the DFS-2 is useful to get a robust picture of a participant's flow without overburdening them (Jackson, Martin, & Eklund, 2008). It consists of eight items and can be aggregated into one single flow score. Participants answered on a Likert-scale ranging from 1 (never) to 5 (all the time). All eight items and their translations can be found in Appendix J.2.

Similarly to the GEQ, not enough participants took part to perform any parametric tests. Once again, I used the Mann-Whitney U test to determine the statistical significance of any difference found (with significance level $\alpha = 0.05$).

6.1.3.3 Knowledge

The filter bubble knowledge test (Appendix J.3) consisted of four items, one of which is an open-ended question (Question 2).

The first question is one primarily for self-assessment of knowledge on filter bubbles. It is not necessarily intended to measure knowledge level, but to compare with students' answers on

[15] Given that there are just three participants in Condition A, these tests can never conceivably yield a statistically significant result. Rather, they were performed as they were intended to be, prior to gathering the data.

Question 2. Additionally, even if pre- and post-activity answers on Question 2 are the same, participants' confidence in their answers could have improved (e.g. from 'I find it hard to explain' to 'I can explain what it is'). The Mann-Whitney U test was used to determine if confidence improvements between conditions differed significantly ($\alpha = 0.05$).

Answers on Question 2 were 'marked' by both myself and Anneleen Janssen, based on the following rubric:

Table 7: Rubric used to mark answers on Question 2.

1	2	3	4
Answer is missing/intentionally left blank/does not relate to filter bubbles at all	Has elements related to filter bubbles, but is not an appropriate explanation of what a filter bubble is	Explains what a filter bubble is, but is missing some (key) elements/is an incomplete picture	Clearly shows knowledge on filter bubbles, their causes and/or consequences

By averaging the two marks per participant per answer, an average score per condition was obtained before and after playing the activity (inter-rater percent agreement = 90%). The statistical significance of the difference in average score was determined using the Mann-Whitney U test (significance level $\alpha = 0.05$).

The third and fourth questions do not have objective truths as 'correct' answers, but are more related to the participants' general awareness of filter bubbles (Q3) and the process of opinion forming (Q4). The activity intends to both raise awareness that participants are most certainly in a filter bubble, and that people can change opinions based on what they see or do online. Given that these questions cannot be objectively graded as being good or wrong, differences in pre- and post-activity answers on these two questions are not statistically tested, but their scores can be compared directly.

6.1.3.4 Qualitative analysis

Recordings of the sessions were transcribed and qualitatively analyzed to retrieve not only general feedback on the activity, but also expressions and statements on the participants' experience. They can be used to support - or contradict - the quantitative results from the knowledge tests and experience and flow scales. Furthermore, the second predetermined talking point is on their experience with managing their own bubble. Contributions made there can be used to directly compare the two (digital vs. tangible) conditions.

Observations during the student sessions concentrated mainly on the use of and interactions with the bubble (either digital or physical) by the participants. Participants forgetting to keep their physical bubble up to date for example could indicate participants having difficulty switching between modalities.

Through open coding, the students' statements and ideas were coded into categories related to the activity or to the usage of the bubble, either digital or physical (see Figure 20 for all final codes).

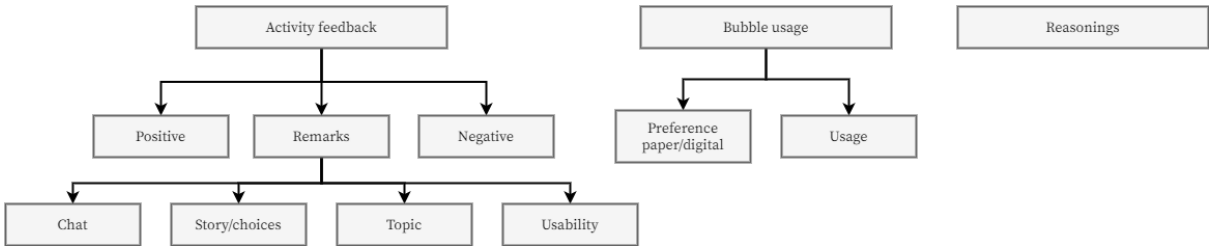

Figure 20: Coding scheme for the evaluation sessions with students.

6.2 Method for teachers

To compensate for the deficit in participant numbers in student evaluations, I also interviewed two teachers who could provide an expert opinion on the distinction between physical and digital methods.

The interviews were semi-structured, and consisted of three distinguishable parts. First, teachers were asked about their experience with paper hand-outs or paper worksheets, both used separately and in combination with digital exercises. Second, they were shown the prototype for Activity 3 and were asked specific questions on what they thought would work best here. Finally, I presented teachers with excerpts and quotes from the student sessions to provoke reactions to those remarks. The full protocol with questions and quotes used can be found in Appendix K.

6.2.1 Participants

The two teachers that participated in this evaluation also participated in earlier phases of this thesis. For consistency purposes, they were given the same pseudonym as in these phases. These specific teachers were recruited not only because of their extensive feedback and expertise, but also because they had already mentioned relevant notions in the previous interviews; for example, D7 reported that he sometimes preferred to use paper sheets in programming classes. Table 8 further details their experience.

Table 8: Participating teachers in the evaluation study. [a]Levensbeschouwing.

ID	Gender	Experience (in years)	Course	Level of education
D1	F	9	Philosophy of life[a]	havo/vwo
D7	M	16	Computer Science, Biology & Head of department havo 4/5	vmbo/havo/vwo

6.2.2 Analysis

After the interviews were transcribed, I marked the key phrases and expressions. I then categorized them into four topics: (1) general use of paper worksheets in class, (2)

(dis)advantages of the paper bubble, (3) (dis)advantages of the digital bubble and (4) general suggestions or remarks about the activity.

6.3 Evaluation results

In this section, I present the results of the final evaluation. It has been split up in two separate sections for the sessions with students and the interviews with teachers.

6.3.1 Students

Data was exported from the online questionnaire and prepared in Excel. Statistics were calculated within Excel, with the exception of the Mann-Whitney U tests, which were performed using SPSS[16].

6.3.1.1 Quantitative results

GEQ. The results of the Mann-Whitney U tests that were conducted on all seven constructs of the GEQ can be found in Table 9. Because it is an ordinal measurement, its medians and interquartile ranges are reported. For all seven constructs, the following hypotheses were tested:

H_0: GEQ score Condition A = GEQ score Condition B
H_1: GEQ score Condition A ≠ GEQ score Condition B

Therefore, two-tailed Mann-Whitney U tests were performed. *P*-values were tested against the Bonferroni-corrected alpha of .007 (0.05/7).

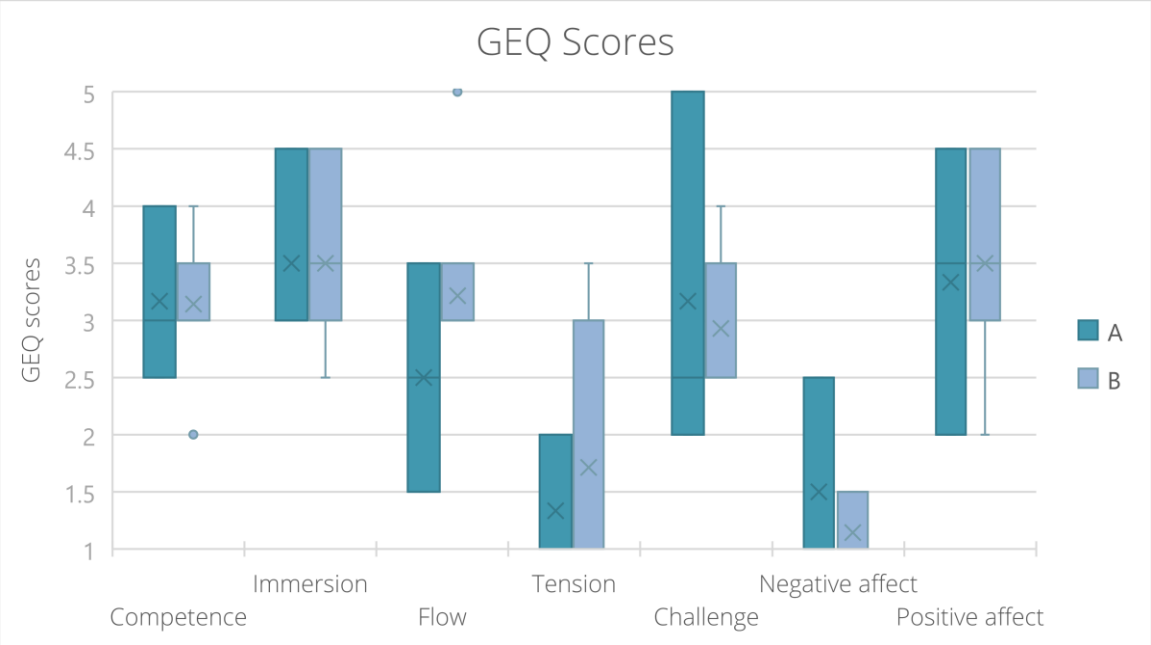

Figure 21: Boxplots for the GEQ scores of both conditions.

[16] https://www.ibm.com/nl-en/analytics/spss-statistics-software

Table 9: Results of the GEQ for both conditions, and p-values of the Mann-Whitney U tests.

GEQ construct	A Mdn (n=3)	A IQR (n=3)	B Mdn (n=7)	B IQR (n=7)	*p*-value
Competence	3	0.75	3	0.5	1
Sensory and Imaginative Immersion	3	0.75	3.5	1	1
Flow	2.5	1	3	0.25	0.409
Tension	1	0.5	1	1.25	0.797
Challenge	2.5	1.5	2.5	0.75	0.808
Negative affect	1	0.75	1	0.25	0.778
Positive affect	3.5	1.25	3.5	1.25	1

None of the Mann-Whitney U tests for the GEQ constructs returned a statistically significant difference between the two conditions, which in all likelihood is due the low number of participants per condition. The most notable difference between condition A ($Mdn = 2.5$, $IQR = 1$) and condition B ($Mdn = 3$, IQR = 0.25) was for the construct of flow, but not significant ($U = 6.5$, $p = 0.409$). Therefore, for all constructs, the null hypothesis could not be rejected.

DFS-2. The results of the Mann-Whitney U tests that were conducted on the general DFS-2 flow score and its eight constructs (Cronbach's alpha = 0.64) can be found in Table 10. Because it is also an ordinal measurement, its medians and interquartile ranges are reported. To calculate the individual DFS-2 scores, the construct scores were added up and divided by the number of constructs (eight), as recommended by Jackson, Martin, & Eklund (2008). For all eight constructs and the overall score, the following hypotheses were tested:

H_0: DFS-2 score Condition A = DFS-2 score Condition B
H_1: DFS-2 score Condition A ≠ DFS-2 score Condition B

Because of this, two-tailed Mann-Whitney U tests were performed against a Bonferroni-corrected significance level of .005 (0.05/9).

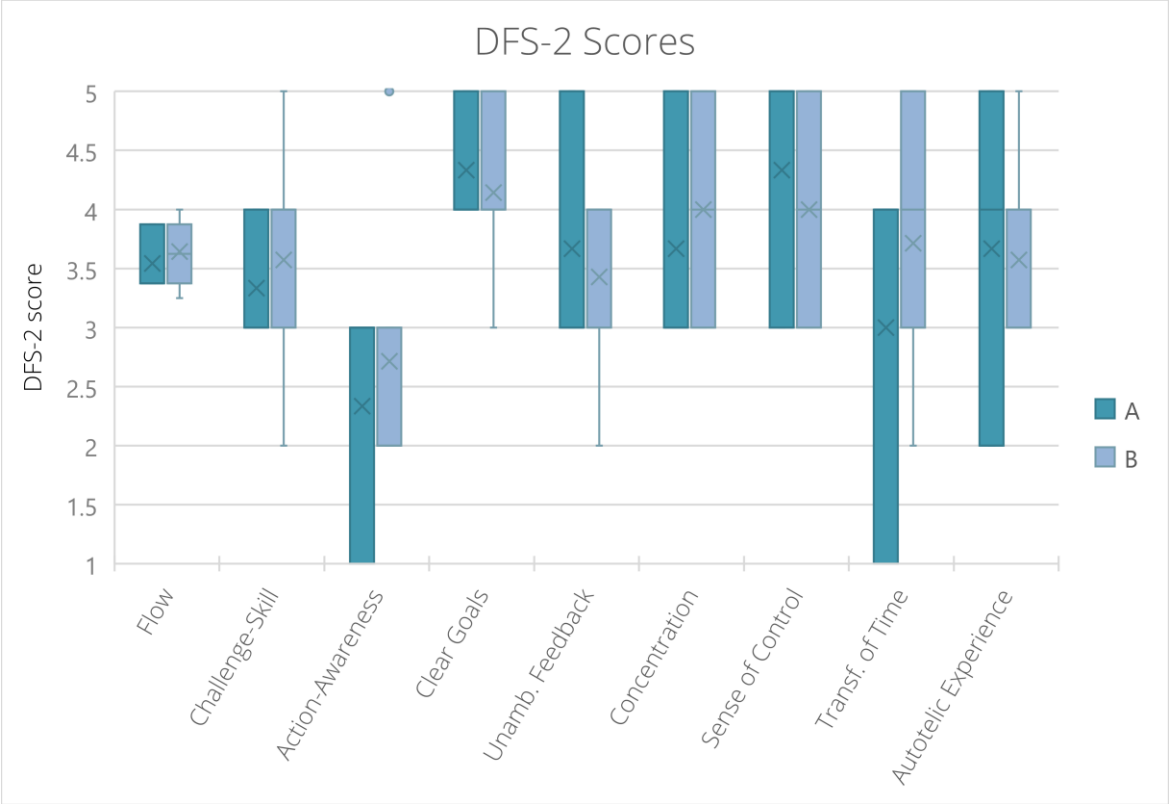

Figure 22: Boxplots of the DFS-2 scores for both conditions.

Table 10: Results of the DFS-2 for both conditions and p-values of the Mann-Whitney U tests.

DFS-2 construct	**A Mdn (n=3)**	**A IQR (n=3)**	**B Mdn (n=7)**	**B IQR (n=7)**	***p*-value**
Flow	3.38	0.25	3.63	0.44	0.640
Challenge-Skill Balance	3	0.5	4	1	0.715
Action-Awareness Merging	3	1	2	1	1
Clear Goals	4	0.5	4	0.5	0.794
Unambiguous Feedback	3	1	4	1	1
Concentration on Task at Hand	3	1	4	1	0.629
Sense of Control	5	1	4	1	0.629
Transformation of Time	4	1.5	4	1.5	0.636
Autotelic Experience	4	1.5	3	1	0.905

While for this sample, the overall DFS-2 flow score was higher for condition B (*Mdn* = 3.63, *IQR* = 0.44) than for participants in condition A (*Mdn* = 3.38, *IQR* = 0.25), this difference was not statistically significant ($U = 8.0$, $p = 0.640$). None of the differences in sub-construct scores of the DFS-2 were statistically significant, and accordingly, none of the null hypotheses could be rejected.

Knowledge. Question 1: Compared to before playing the activity, 7 out of 10 students said their knowledge on filter bubbles improved (3/3 in condition A, 4/7 in condition B). The other three students (all condition B) indicated their knowledge did not improve (see Table 11).

Table 11: Scores on Question 1 on self-assessment of knowledge on filter bubbles for both conditions. 1 = 'No, and I've never heard of it', 2 = 'No, but I have heard of it', 3 = 'Yes, but I find it hard to explain', 4 = 'Yes, and I can explain what it is'

Condition	Participant	Before	After	Difference
A	P10	1	2	**+1**
(physical)	P11	2	3	**+1**
	P12	3	4	**+1**
B	P5	1	4	**+3**
(digital)	P6	3	4	**+1**
	P7	1	2	**+1**
	P8	2	2	0
	P9	1	1	0
	P13	2	3	**+1**
	P15	3	3	0

The Mann-Whitney U test (with H_0: improvement in condition A = improvement in condition B, H_1: improvement in condition A ≠ improvement in condition B) did not reveal any statistically significant difference ($U = 13.5$, $p = 0.514$) between condition A (*Mdn* = 1, *IQR* = 0) and condition B (*Mdn* = 1, *IQR* = 1). Since $p > \alpha$, H_0 could not be rejected.

Question 2: Two participants (P11 in condition A, P5 in condition B) improved in their score on Question 2 after the activity compared to before the activity. They both scored a 1 prior to the activity, and received an average of 1.5 for their answer after the activity (see Table 12). These participants were both in the group of seven participants that indicated in Question 1 that their knowledge of filter bubbles had improved. All eight other participants scored a 1 both before and after the condition: seven wrote down 'I don't know' (or a comparable phrase), one gave a wrong answer. Since scores only improved for two participants, both in a different condition, no significance test was performed to determine statistical significance of any differences, given that there clearly was no difference.

Table 12: Scores on Question 2 on explaining what a filter bubbles is, for both conditions. The rubric that was used to grade answers can be found in Section 6.1.3.3.

Condition	Participant	Before	After	Difference
A	P10	1	1	0
(physical)	P11	1	1.5	**+0.5**
	P12	1	1	0
B	P5	1	1.5	**+0.5**
(digital)	P6	1	1	0
	P7	1	1	0
	P8	1	1	0
	P9	1	1	0
	P13	1	1	0
	P15	1	1	0

Question 3: Before the activity, 7 out of 10 participants did not agree they were in a filter bubble themselves (answers 1 or 2), of which 3/3 in condition A and 4/7 in condition B (see Table 13). The other three participants responded neutrally (answer 3). No participant agreed they were in a filter bubble (answers 4 or 5). After the activity, five participants changed their answers. Three participants now agreed to the statement that they were in a bubble themselves (1 in condition A, 2 in condition B), while two other participants disagreed more compared to before the activity (1 in condition A, 1 in condition B).

Table 13: Scores on Question 3, on if students agreed they were in a filter bubble themselves.

Condition	Participant	Before	After	Difference
A	P10	1	1	0
(physical)	P11	1	5	**+4**
	P12	2	1	**-1**
B	P5	1	5	**+4**
(digital)	P6	3	5	**+2**
	P7	1	1	0
	P8	1	1	0
	P9	1	1	0
	P13	3	3	0
	P15	3	1	**-2**

Question 4: Before the activity, 7 out of 10 participants indicated that they agreed (answers 4 or 5) with the statement *What people see or do online can change their opinion*, of which 2 in condition A and 5 in condition B (see Table 14). After the activity, two participants (both condition A) changed their answer: one lowered their agreement to neutrality (answer 4 to 3), one went from neutrality to agreement (answer 3 to 4).

Table 14: Scores on Question 4, on if students agreed online behaviour can influence opinions.

Condition	Participant	Before	After	Difference
A	P10	3	4	**+1**
(physical)	P11	5	5	0
	P12	4	3	**-1**
B	P5	5	5	0
(digital)	P6	5	5	0
	P7	3	3	0
	P8	3	3	0
	P9	5	5	0
	P13	5	5	0
	P15	5	5	0

6.3.1.2 Qualitative results

In condition B, with the digital bubbles embedded in the application, three out of seven participants did not properly keep their bubble up to date. Two indicated they had forgotten to do it during the activity, while one participant misunderstood where they could find the button to open the screen with their bubble. The other four participants used their bubble for every choice they made. In condition A, with the paper bubble, one participant did not immediately understand what to do with the paper choice cut-outs: "*But I already made this choice on-screen?"* (P12). During the activity however, all three participants used their paper bubble for every choice they made.

After participants showed their bubbles, they had made during the activity to other participants who had played with another character, they demonstrated that they understood why this was the case:

> *"Because she subscribes to certain* [YouTube-]*channels for example. She then gets to certain videos, other videos than Jayden."* (P5)

This reasoning did not differ between the conditions. When asked if such a separation of opinion through online behaviour could also happen in real life, almost all students agreed, but none could provide other examples, even if they indicated this had happened to themselves as well.

Finally, students were directly asked if they would prefer to have the paper bubble or the digital bubble. Irrespective of the condition they were in, they all preferred to have the bubble embedded within the application. They gave a multitude of reasons: participants in one session (G1) all worried that quarrels would ensue if paper was used:

> *"Using paper, it's really going to be… They'll* [other students] *start pulling on the papers, "I want to do it, I want to do it", quarrelling, the teacher would then take away the paper, and then we all can't play anymore. But on the laptops, nobody would do that."* (P5)

Others, such as the participants in G2, thought the digital bubble to be more engaging. P15 from G3 also supported that:

> *"Online, because* [other] *students are on their phones a lot in general, so I think it would be nice if you could bring something like this directly to them, online. Because that would appeal a bit more. Or make it more fun."* (P15)

Participants also named some disadvantages of working with paper: it would distract them (*"often when using paper, everyone is a bit more quickly distracted"*, P13), thought it to be less clear ("*using paper it's less fun and it won't be entirely visible"*, P8) and P15 also named the large amount of paper needed as a disadvantage.

6.3.2 Teachers

Before being shown the activity, teachers reported on their own experiences with using paper sheets in class. D1 mentioned that combining digital exercises with physical worksheets often turned out chaotic and cluttered, causing her to often abandon digital exercises altogether. D7, being a computer science teacher, said about the digital versus physical divide that "*you have to employ techniques when they're valuable* […] *and I don't think that digital is always better"*. They for example mentioned that when there is a need for automatic feedback or scores, digital would be better, but for cooperation or moments where computers could distract, working on physical paper would be better.

After being shown the prototype of Activity 3, both teachers had a clear preference for the paper implementation of the bubble for multiple reasons, although D1 also indicated this would be dependent on personal preference as well. Primarily, both teachers immediately indicated a paper sheet would be far more accessible and insightful for them as teachers:

> *"If there's a big sheet of paper like that on the table, and I want to say something about it or point out an error to them or something like that, I just have to point at it. You can just literally make it tangible. And otherwise, you're on this little screen, so I still have to actively go to the students and get very close to them to be able to point out to them what stands out to me.* […] *And watching along on a screen, you'd have to get really close* […] *and that's pretty invasive."* (D1)

> "*Then, I think this would be very helpful on paper, because as a teacher, you'd be able to literally see on the table what decisions they* [students] *made*." (D7)

While both teachers agreed that in the case of digital bubbles, a teacher dashboard to provide an overview in their students' bubbles would be helpful, they would still prefer the paper version:

> "*Well, that would make digital* [bubbles] *a bit easier, but I think as a teacher, with this kind of an assignment, you're just walking around the classroom all the time, motivating kids and giving them a little extra food for thought, or a little push in the right direction. You're not sitting at the front of the classroom, looking at your dashboard, while the children are busy."* (D7)

D1 also saw the same advantage for students: to discuss the differences in their bubbles, they would not have to "*carry around their laptops everywhere, and just show the papers".* Likewise, D7 anticipated that the paper bubble would be better visible for other classmates than the digital implementation, which would be advantageous during the activity as well:

> "*Because around you, you'd see other students are also putting things in the circle. So like 'oh, they already got two things in the circle, I haven't got anything yet, I must have forgotten something'."* (D7)

Because of the visibility of the paper bubble, D7 predicted that this version would have the most educational value as well. D1 doubted that there would be much of a difference between the two versions in this respect. D1 also indicated that students in her class would not necessarily enjoy one version much more than the other, but did say that working with tangible objects with their hands would be appreciated by some students, which D7 supported:

> "*Yeah, some* [students] *like to do something practical for a while, like cutting and pasting. And of course that's not there if you do dragging and dropping, because that's quickly over."* (D1)

> *"And it's just nice to physically shuffle those things around, to do something with your hands."* (D7)

Teachers had dissimilar views on the possible disruption of flow that paper bubbles could cause. D1 thought the digital bubbles would allow students to stay more focused, while D7 though that that 'disruption' of flow would actually be beneficial:

> *"Then maybe digital would be better [for students' concentration]. Because there's less of a possibility of noise and hassles. There's no papers falling on the floor, no arguing about who's going to cut the papers… So then, I can imagine that digital [bubbles] is more efficient."* (D1)

> "*I think it's a good intermediary moment* [using the paper bubble], *instead of 'I'm racing through a digital program', that it's like 'oh right, I'm one step further, I can physically move something'. So instead of just, click, progress, that there is a physical moment to proceed. I can imagine that then they'd do things more consciously, and therefore learn more than they would*

> *if everything goes automatically, if it's just clicking through and collecting things. I think there's more value in doing that physically."* (D7)

Both teachers could relate to the students who said during the discussions that using paper bubbles could see quarrels and practical issues emerge in class (*"in that respect, a digital version is more ready-made",* D1). D7 however also commented that if a teacher would prepare their lesson well, it would not be as bad as the students made it out to be - although D7 did agree that such a lesson would take a lot of preparation.

Conclusively, teachers preferred the version of the activity with the paper bubble, as articulated by D7:

> *"So I think this is actually a very nice hybrid format, that you go through those steps digitally, and that this [paper bubble] actually is the 'journal' of the choices you've made."* (D7)

6.4 Evaluation discussion

In this evaluation study, I aimed to evaluate the two possible implementations of Activity 3, and discover whether the presence of a physical artefact improves (1) students' experience of the activity, and (2) their knowledge and awareness of filter bubbles. Because of the low number of participating students, the quantitative data proved to be inconclusive, but on the basis of observations and qualitative input, students have a clear preference for the digital implementation. However, teachers would rather see the paper bubble.

Regarding the GEQ (Game Experience Questionnaire), which was intended to measure differences in experience between the two conditions, there were no statistically significant differences between the experiences with a physical or digital bubble. The DFS-2 also showed no different scores for participants in different conditions, and much like the GEQ results have no ground for any definitive conclusions to be drawn. For both conditions, the GEQ did show low scores on *Tension* and *Negative affect*, which is favourable for the prototype as a whole as students did not feel bored or frustrated with the activity.

The knowledge test provided some interesting results. First of all, in both conditions, students indicated in Q1 that their knowledge of filter bubbles had improved after playing the activity but did not improve their answers on Q2 which actually tests that knowledge. This could indicate that students' confidence in their knowledge increased after playing the activity (irrespective of using the paper or digital bubble), but their actual knowledge did not. Another possibility is that the phrasing of the question itself was inadequate. In any case, the absence of improvement in Q2 further underline the importance of a teacher-guided class discussion after the activity, and that the activity should not be used as a stand-alone lesson.

The low scores in Q2 also impact the results of Q3, as students need to know what a filter bubble is before they can indicate if they themselves are in a filter bubble. The possible answers students could give on this question did not include a 'I don't know' option, which means that it cannot be determined whether the responses students gave were genuinely indicative of them thinking they are in a filter bubble, or not knowing what a filter bubble is. As for Q4, before the

activity, students already indicated they thought online behaviour could influence opinions and attitudes and did not change their answer after the activity. This was also reflected in students' reasoning about why Jayden and Sophie ended up with different videos and topics in their bubbles, which showed that they are well aware of them seeing and coming across different things on the internet than others. The concept of a 'filter bubble' however does not yet seem connected to that belief.

While the quantitative data did not show an advantage for one condition above the other, the short discussion after playing the activity revealed that all students would prefer to have the digital bubble instead of the paper bubble. Their reasons for this preference ranged from perceived fun and engagement, practical issues and environmental objections, in accordance with Miglino et al. (2014) and Liao et al. (2007). On the other hand, teachers preferred the physical paper implementation, primarily because that would be more useful and convenient for them in class. In line with previous literature, one teacher predicted that that continuous switch between physical and digital would come at the cost of less concentration. The other teacher however took another approach and said that this 'disruption' of flow would be beneficial, as to not let students rush through the application but needing them to be more conscious and mindful of keeping their paper bubbles up to date.

These results build on the existing evidence supporting and opposing both the physical and digital possibilities. There is something to be said for both implementations. The paper implementation is preferred by the teachers, makes collaboration easier, and has numerous other practical advantages, but also comes with a significant ecological footprint, more burden on the teachers to prepare their lessons and students seemed more enthusiastic about the digital bubble. So while the presence of a physical artefact does not seem to improve students' experience of the filter bubble activity (RQ1), according to the teachers it does improve the educational value and therefore the knowledge and awareness of students on filter bubbles (RQ2). Because of this convincing preference of these two teachers, the paper bubble seems the most appropriate implementation. Teachers need to be comfortable with their instructional materials and are ultimately the ones who decide if a lesson is going to be successful or not. Because of this, based on the results of this study, I recommend Activity 3 to be played with the paper version of the bubble.

These results also should be taken into account in future HCI developments for education. Teachers evidently do not necessarily want educational applications to take place solely in the digital domain, and this should be considered when developing educational tools; there is an opportunity for physical, tangible 'add-ons' to enrich activities or exercises. As demonstrated by this study, the advantages of such physical additions may outweigh the disadvantages by facilitating discussion and collaboration.

6.4.1 Limitations and future work

The generalizability of this evaluation study is limited by a number of factors. First and foremost, the small sample size combined with the unequal distribution of participants among the two conditions made it impossible to infer any generalizable conclusions from the quantitative data. Additionally, one of the sessions (G3) was held online, and while this did not change how

students did the activity, it still might have impacted the validity of results as it was not held in the same setting as the other two sessions. Another limitation is that some participants had already seen a prototype of another activity in some shape or form, while others had not, and that these participants could not be spread evenly among conditions. As a result, all participants in condition B had already participated earlier in the design phase, while participants in condition A had not.

While the lack of student participants was somewhat made up for by the inclusion of two teacher participants, their results also suffer from limited generalizability, given that only two teachers took part. D7 for example had already indicated their personal preference for using paper in his computer science classes prior to the evaluation, possibly resulting in a bias towards the paper bubble implementation.

The validity of the quantitative data is impacted by the lack of properly translated Dutch versions of both the DFS-2 and the GEQ. The DFS-2 was not at all available in Dutch, so I had to translate it myself. Even though the translations are as close as possible to the original English versions, these Dutch questions have not been validated, and the Cronbach's alpha level of 0.64 indicates a questionable internal consistency. The GEQ was available in Dutch and for the most part directly reproduced, but a number of items (indicated in Appendix J.1) were altered to accommodate for the language level of vmbo students.

In future research, other quantitative measures could also feasibly be used. In a larger scale study, the PANAS (*Positive And Negative Affect Schedule:* Watson & Clark, 1999) could be used to more reliably measure students' positive and negative affect during both conditions. For the current study, PANAS was deemed unfeasible because of its length of 60 items. Alternatively, the SUS (*System Usability Scale*: Brooke, 1996) could be used to test differences in perceived usability of the two conditions.

To properly validate the educational value of the activity and of both conditions, the activity should be evaluated in a natural class setting with actual teachers guiding full classes through the lesson, as the activity was intended. Such a complete lesson could provide further insight in how the paper bubble does or does not facilitate the class discussion as suggested by the teachers, and how students interact with the activity when not in the small focus group-like setting. Should the digital implementation demonstrably improve flow, but the paper bubble yield better results learning-wise, this could cast doubts over the notion that flow is crucial for engagement with game-based learning environments, as hypothesized by Admiraal et al. (2011).

6.4.2 Evaluation conclusion

Previous studies have pointed to the benefit of having tangible, physical objects to move around during learning. However, having to interact with a physical object while going through a digital activity might disrupt the state of flow of students, which has adverse effects on engagement and ultimately learning. In this evaluation study, I attempted to determine the influence of having a physical artefact on students' experience of and knowledge gaining in Activity 3 and contribute to the literature on this dichotomy between tangible and fully digital experiences. Three sessions with in total ten students (three with paper, seven with digital bubble) were conducted to gather

both quantitative and qualitative input on their experience, flow, knowledge gains and personal preferences. Additionally, two teachers were interviewed to provide their expertise on both teaching methodologies.

Although quantitative results were inconclusive because of the distribution of participants, qualitatively the sessions with students and interviews with teachers provided valuable findings. Students preferred the digital version of the bubble, while teachers were adamant that the paper bubble would work best in their classes, and all had valid and well thought through reasons for their preferences. Ultimately, although personal preference definitely plays a role, the arguments teachers gave for opting for the tangible version were the deciding factors in recommending Activity 3 to be played with the paper bubble. Future work should evaluate both possibilities in a real class setting to confirm these findings.

7. DISCUSSION

Throughout this entire study, by testing and redeveloping the prototypes of activities with students and teacher, I have attempted to answer the following research question:

RQ *What are the design qualities of interactive tools that effectively raise awareness among teenagers?*

In this chapter, I present the recommendations for both future development of the application and for HCI with teenagers that have emerged as answers to this research question. This chapter also contains the limitations of the study and the conclusion.

7.1 Recommendations for future development

Below, I present the itemized recommendations for future development of the application in general (Section 7.1.1) which also apply to the broader context of developing educational applications for teenagers. I also present recommendations for the two specific activities (7.1.2 and 7.1.3), and for future HCI research with teenagers (7.1.4).

Regarding the overall idea (see Section 5.5), it is vital that - after further development - the idea is evaluated with students, as this has not been done yet. Preferably, this should be done in combination with one of the activities, as to simulate an entire lesson.

7.1.1 General recommendations

While these general recommendations can be directly applied to the anti-filter bubble application, they also serve as general recommendations for any similar HCI project in secondary education.

Empower the teacher. Building on Bower (2017) and Kolb (2017), the role of the teacher is even more important than previously thought. The requirements already stated that the teacher should have a clear overview of what their students are doing in class with the application (Req 2) and that the application should help the teacher in fostering a safe environment (Req 3). It has now become clear that the teacher is key to a positive experience for students and to ensure that students actually learn something. Students will learn during the discussion that follows the activity in the application, not necessarily during the activity. The activity serves as a steppingstone towards the class discussion. This is something that Kolb (2017) also mentioned - digital educational tools should not just be focused on engagement, but also enhancement (in this case, the activity enhancing the class discussion) and extension (in this case, the activity providing real world examples). It is important to realize this when developing new activities - they should not be intended to last longer than five or ten minutes unless some kind of class discussion is already embedded in the activity. Teachers should be given a lot of freedom and space to be able to host class discussions as they see fit, and should not be forced to adhere to a predetermined lesson plan. Within activities, teachers should also be given the opportunity to make choices about certain aspects, such as including collaboration or having students do the activity individually.

Provide a teacher's guide. All teachers, regardless of their experience or subject, indicated they would like to see some kind of manual, guide, handbook or lesson plan. Such a guide serves many goals: first and foremost, it introduces teachers to the application and guides them through the steps they need to take to get it to work in their class. Additionally, it should provide less experienced teachers with ideas for how to handle class discussions, and should contain more content-related information to introduce teachers who lack the relevant knowledge to filter bubbles, algorithms, and other important topics. Connecting the goals of the application to core educational objectives or final attainment levels can help teachers place the application in their plans for the year. Finally, it is important that this guide contains suggestions, and not obligations, as teachers should be able to shape the lesson as they see fit.

Make sure the application supports a safe atmosphere. Already mentioned in Requirement 3, the application should help foster a safe environment for students. Under no circumstances should the application or its activities undermine the atmosphere in class. This is a very important prerequisite for many teachers. The above-mentioned teacher's guide can help teachers in this, but it is also important that the activities and the application do not force students to do things or say things that they do not feel comfortable with. An example of this is the now discontinued Activity 1 (Section 5.2), which would have obliged students to share their personal YouTube recommendations with the entire class. The freedom for the teacher to organize the lesson the way they want it to also provides them with the opportunity to adapt activities to specific classes, such as choosing to not use collaboration in a class where that is difficult to arrange.

Decide on the application's place in the curriculum. While this is undoubtedly depending on schools themselves, attention needs to be paid to the place this application will have in schools' curricula. Seven lessons take up multiple weeks in an already crammed schedule, which is also why multiple teachers responded hesitantly when told the entire application would take seven lessons to use completely. Solutions could be to create either an alternative, shorter version of the application that spans three or four lessons, or to make sure that activities can also be used separately from the application as stand-alone lessons. Nonetheless, the further development of the application should stay in close communication with the participating project schools as to make sure that there is a consensus on the place of the application within their curricula.

Ensure content is easily replaceable. The content within activities, such as videos in Activity 2 or choices in Activity 3, should be easily replaceable. This is important for two reasons: first of all, I have noticed the importance of relevant content for students. Already mentioned in Requirement 5, students are immediately stimulated when being shown relatable content, such as YouTubers they are familiar with (e.g. Touzani in Activity 2) or topics that they have a strong opinion about (e.g. family vlogging in Activity 3). These trends change rather swiftly, and while bigger themes such as children in family videos will stay relevant for some time, specific videos and personalities may not. Secondly, specific content needs to be able to be dropped quickly in case outside events make it no longer appropriate to be used in an educational setting. This does mean that there needs to be a dedicated development team or individual updating the application, or at least monitoring the need for updates, every few months.

Test and improve usability and UX. The work that has gone into the prototypes thus far is primarily content-based, and while work has gone into usability and minor usability improvements based

on feedback have been implemented, it needs dedicated usability evaluations and developments, as these were outside the scope for this thesis. Prototypes were intentionally kept in the same colourless, minimalistic style, as to minimize the impact of UI design on students' impressions, but is obviously not resembling the intended style of the final product. The user experience as a whole needs to be tested in multiple user tests, as this is a core aspect determining a student's enjoyment of the application. Students need the application and its activities to look and feel good. Additionally, the current prototypes were developed for desktop or laptop screen sizes. Given it is intended to be a responsive web-application, it should function on mobile phone screens as well, which means that the lay-out (and therefore the UX) on smaller screens needs to be considered as well. Finally, no work has yet gone into accessibility of the application for for example colour-blind students or the visually impaired. While this recommendation is relatively trivial, it is crucial that the need for UX improvements and usability tests is not underestimated - this can and will make or break activities and the application as a whole.

Evaluate the activities in classrooms. All the prototypes were tested and played with students in focus groups. While these focus groups took place at their schools, it is far removed from an actual class setting with a teacher guiding them through the activity and 25 other students being present. In order to properly judge the effectiveness and students' enjoyment of the activities, they should be tested during a real lesson.

7.1.2 Activity 2

Activity 2 was received very positively by both students and teachers, and has gone through several iterations. However, there are some matters that need attention before further development can take place.

Algorithm table. The algorithm table that was present in the focus groups, on which students could track what Robin did and did not like, has not been implemented within the digital prototype itself but rather is a physical artefact complementing a digital activity. Put this way, it bears resemblance to the dichotomy present in the evaluation of Activity 3. Based on those results, I would recommend the algorithm table to be used on paper as well, but given that the type of interaction with the algorithm table is different than with Activity 3's bubble (typing text vs. dragging items), the same conclusion cannot be drawn. However, the version with the paper algorithm table already has been used and tested, while a digital implementation would still need to be developed and tested again. For that reason, retaining the use of the paper algorithm table seems most feasible.

Different topics, viewers or storylines. Robin's storyline, with his subscriptions to Milan Knol, Touzani and Supergaande, is one of many that could be implemented in the final activity. It is possible to replace the videos or topics for the sake of topicality, as mentioned in 7.1.1, or let teachers choose between multiple storylines. It is important that other storylines follow a similar setup - the full list of videos in the current storyline can be found in Appendix H - as to not have large difficulty gaps between different storylines.

Second round. Multiple teachers suggested adding a second round to the activity, which would contain either a more polarizing storyline where Robin is recommended increasingly more extreme videos, or videos that are less based on entertainment but more political. According to them, this would help make the underlying premise and danger of filter bubbles more explicit. While certainly more straightforward, it would also introduce a risk of unearthing controversial topics in class, possibly thwarting the safe atmosphere in class. Should such a second round be introduced therefore, it should not be mandatory or required in order to progress, but serve as a supplementary activity if the teacher chooses to do so. Another viable possibility is that a second round is more technically focused, and teaches students about distinct types of recommendation algorithms, which is also recommended by RAN (2018) to improve digital literacy.

Collaboration. From the feedback received in the design phase, no conclusions could be made on whether to have students work in groups during this activity. Anneleen Janssen has evaluated the activity in both groups and individual settings, so for an analysis on this topic, I refer to her thesis.

7.1.3 Activity 3

Activity 3 teaches students about taking perspectives and understanding others' opinions, while also introducing students to the effects that online behaviour can have on opinion forming and introduces the filter bubble to them. While having gone through relatively few iterations, the activity was generally received well among both teachers and students.

Digital or physical bubble. The final evaluation (Chapter 6) was aimed at finding out whether a physical, tangible implementation of the bubble in Activity 3 would result in better experience and knowledge gains than a digital bubble. It turned out that while the students all preferred a digital implementation, teachers clearly saw more value in the paper bubble. Therefore, I recommend the paper bubble to be used with this activity. Further details on the (dis)advantages of both possibilities are presented in Chapter 6. An additional final improvement that needs to be made is to add other items to the 'outside' of the bubble as well (e.g. Jayden's choices in Sophie's storyline), for students to be able to see the big picture of items that they have not come across during their story.

Storylines. The storyline in the prototype, about children in family vlogs, has two developed characters (Jayden and Sophie). It was intended to have four characters (Younes and Kyra as well), but for practical purposes was limited to just Jayden and Sophie in focus groups. This means that these stories of Younes and Kyra should still be developed and implemented in the prototype. Additionally, all four stories should be made somewhat longer, as students went through it rather quickly and indicated themselves that they would like a few more choices. Another possibility is to develop more storylines on different topics, and have the teachers - or even students - choose what topic is most suitable. While this is not a necessity for this activity, it might accommodate classes that are less interested in or less familiar with family vlogging.

Chat. The second part of the activity, in which students that played with different characters chat with each other, has only been shown twice to students and was not playtested like the rest of

the activity, although it was received very well among those who it was shown to. However, the chat assumed the bubble would be implemented in the digital prototype, and consequently, would have students drag their bubble content to the chat to indicate where they got certain information from. With a paper bubble, this feature disappears, and the added value of the chat becomes nil; students would directly discuss the contents of their tangible bubbles with each other in class. Further development and testing of the chat thus depends on whether the paper bubble is implemented.

Replayability. The intention of the activity was to give students insight into the consequences of their online choices by letting them compare bubbles with each other. Teachers indicated that providing the possibility to replay the activity with different characters or to make retake choices could also be used to attain this goal. The opportunity to retake choices was also met with enthusiasm from students when suggested during focus groups. If this would be implemented, I would recommend providing students the opportunity to do so after they discussed the differences in bubbles with classmates, to not defeat the purpose of playing with different characters.

7.1.4 Methods in HCI with teenagers

Given the underrepresentation of research for and with teenagers in human-computer interaction, as addressed by Rose et al. (2018), I present several recommendations for future HCI research with teenagers in this section. They are based on the focus groups that were held during this thesis.

Group composition. In line with previous research (such as Fitton et al. (2016) and Raby (2010)), the way a group is composed has considerable impact on how well the focus group runs. If students are friends with each other or at least friendly to one another, they get more comfortable, are encouraged to say what they want to say and dare to call things others say into question. This goes both ways; if students do not necessarily feel friendly towards one another, they clam up. To illustrate, in the pilot focus group for the final evaluation, three students were obviously friends, but one was not part of that group, leading them to say almost nothing during the focus group. Ideally, teachers should be consulted beforehand to ensure the compatibility of participating students. Regarding group size, focus groups in this project differed in size between two to five students. Both small and larger focus groups have their (dis)advantages: smaller focus groups allow for more personal questioning and elaboration, while larger focus groups have more opportunities for reactions and discussions among participants.

Participant selection. For most focus groups, teachers selected students they thought would be best in articulating their opinions and feedback about the prototypes to us. While it could be argued that this is not representative of the entire student population, I would argue that especially in this population of younger students, it is at least better than randomly selecting students that may not be able to offer any useful feedback. In some cases, it might be beneficial to include adults that students trust in the focus group, like their teacher or a youth worker. Their presence helps shy students to speak out and give their opinion. In advance, agreements must be made with that adult about what they can and cannot say to prevent them from polluting the data.

Atmosphere. As mentioned in literature (e.g. Poole & Peyton, 2013), the skewed power relationship students assume when in a focus group can prevent them from speaking freely, and to prevent this, the focus group should be kept as informal as possible. Students playing the activity and being engaged with it was already a big step towards them enjoying the focus group. When students noticed we were also genuinely interested in their personal opinions and interests on social media, they opened up more. Additionally, supporting for example Kinnula & Iivari (2019), convincing students that their feedback is going to be used to improve the prototypes and showing them changes made based on feedback from other students showed them that they have a genuine impact on the design process, and further increased their motivation. The lack of extrinsic rewards but the willingness of participants to continue and participate in future focus groups shows the intrinsic motivation that arose from these factors.

Material. While it might be tempting to show and let students interact with early prototypes in the beginning of the design process, prototypes do need to be foolproof and near infallible before students can go through them on their own. Usability issues need to be ironed out as much as possible and I advise to ensure that everything that is shown in the prototype is also functional, to avoid confusion among students. Additionally, in accordance with Bell (2007), what students can understand language-wise (in for example questionnaires) is hard to predict in advance, and pilot tests are needed to identify any words or phrases that are too difficult or unclear. Using external sources, such as *ishetB1*[17], or involving experts, such as those from Mira Media, is recommended to filter out as many occurrences of these as possible in advance.

Conclusively, these recommendations are based on the focus groups that were held for this project and this thesis. It is important to realize that the participating students were also a very specific group of teenagers (vmbo, 12 to 14-year-old), and as Fitton et al. (2016) stressed, methodologies with teenagers should be adapted to specific contexts and there is no 'one size fits all' technique. However, given the scarcity of previous research, these recommendations do contribute methodological considerations to the often overlooked field of teen-computer interaction.

7.2 Limitations

There are several limitations applying to this thesis, which I will discuss below. Limitations that are restricted to the evaluation are discussed in Section 6.4.1.

First of all, all focus groups were held at schools in Utrecht and Amsterdam. Even though the students that participated had a wide variety of cultural backgrounds, these are two of the most populous cities in the Netherlands. Additionally, just four schools in Utrecht were involved, and even between these schools themselves differences in pre-existing knowledge or attitudes were found. Given that the application is primarily intended to be deployed in the participating schools in Utrecht, this is not an immediate issue, but further evaluations in other areas of the Netherlands have to be conducted before it can be deployed elsewhere. As mentioned before,

[17] http://www.ishetb1.nl

these evaluations should be done in an actual class setting, as the prototypes have thus far only been used in a focus group setting which is not representative of that of a full class.

Throughout the pre-studies and design phase, the teachers that participated were selected either because they teach or taught on vmbo or because they teach or taught courses in which the application could conceivably be used (such as civics or computer science). This meant that we primarily spoke to teachers who teach civics, and those are more experienced in - for example - tackling sensitive topics like polarisation than others. This therefore somewhat limits the generalizability of the results of the teacher interviews.

The covid-19 pandemic, aside from critically restricting participant recruiting, also produced some limitations to the current study. Some focus groups with students were forced to be conducted online, which reduced the interaction students could have with the prototype, limited our possibilities for observation and decreased the overall quality of the focus groups.

7.3 Conclusion

By employing an iterative user-centred design approach and in close collaboration with other project members, I have developed three prototypes for activities and an overall structure to be used in the anti-filter bubble application. The anti-filter bubble application is intended to be used in vmbo schools in the Netherlands, to increase students' awareness of filter bubbles. The application provides support for facilitating class discussions which teach teenagers the skills and competencies they need in a digitally saturated world. I also have contributed recommendations for future development of the application and for methods in HCI with teenagers.

In the first activity developed in this thesis, students take on the role of the YouTube-algorithm and recommend real videos to a fictional viewer. They compete with other (groups of) students to see who can attract and sustain that viewer's attention the best. By doing this activity, students are introduced to the inner workings of the YouTube recommendation algorithm and come to understand why their own YouTube recommends certain videos to them.

The other activity is aimed more at understanding what filter bubbles are and how they form. Students take on the role of one of four fictional characters and play through an interactive story, in which their character is convinced by the videos, news articles and search results they come across that the appearance of children in family vlogs is or is not acceptable. After going through the story, students talk to classmates who have played with another character to find out why their characters' opinions differ. A class discussion then relates these findings to students' own lives as they come to realize that their online behaviour can influence their opinions and what the potential consequences of these filter bubbles can be.

The overall structure for the application that was developed lets students create their own social media platform. In the course of seven separate lessons, students make choices for their own platform related to the content of that lesson's activity. Gradually, after making all choices, their platform is completed. This overall structure helps to add a necessity for learning for students

by means of a large end goal, provides opportunities for personalisation and gives students a sense of ownership.

Furthermore, I also evaluated the influence of the presence of a physical artefact in Activity 3 on the knowledge gains of students and their experience of that specific activity. While quantitative measures on knowledge and experience were inconclusive, qualitative results showed that while students preferred a digital implementation, teachers would rather have the physical paper artefact for multiple reasons. Therefore, I recommended Activity 3 to be performed with the paper bubble.

This thesis provides a first step towards the anti-filter bubble application, and directly contributes to the sparse previous research in human-computer interaction with teenagers. It hopefully preludes further growth of attention to the field, so that teenagers are no longer seen as small adults or tall children, but as genuine stakeholders with their own sets of needs, attributes, and skills.

ACKNOWLEDGEMENTS

I want to thank my supervisor Paweł Woźniak for his weekly feedback, motivation, and support during this thesis. You taught me to trust the design process, how to improve my writing, and above all helped me to always *keep pushing*!

I would also like to thank the pedagogy team members Maartje van Dijken, Mariëtte de Haan, Zena Bani and Anouk Adriani for helping with the focus groups, providing feedback on the prototypes and ideas, and thinking along with us. I am especially grateful to Zena for going out of her way to help us recruit participants in the final stage of the project.

Additionally, many thanks to Judith Masthoff for somehow finding the time and kindness to be my second supervisor and providing feedback. I really appreciate it.

Thanks to Anneleen Janssen for taking up this thesis project with me. During these times of social isolation, it helped to have daily contact and definitely improved my work. Thank you for getting me through it all.

I would also like to thank all teachers and all students that offered up their free time to participate in my study. It is greatly appreciated.

Finally, I want to thank my parents, friends, Lara, and Bart & Nicole for keeping me sane during this project and providing much needed support.

REFERENCES

1. Admiraal, W., Huizenga, J., Akkerman, S., & Ten Dam, G. (2011). The concept of flow in collaborative game-based learning. *Computers in Human Behavior, 27*(3), 1185–1194. https://doi.org/10.1016/j.chb.2010.12.013

2. Amrollahi, A. (2019, December). *Burst the Filter Bubble: Towards an Integrated Tool.* 30th Australasian Conference on Information Systems, Perth, Western Australia.

3. Antle, A. N., Droumeva, M., & Ha, D. (2009, June). Hands on what? Comparing children's mouse-based and tangible-based interaction. In P. Paolini & F. Garzotto (Eds.), *Proceedings of IDC 2009 - The 8th International Conference on Interaction Design and Children* (pp. 80–88). ACM. https://doi.org/10.1145/1551788.

4. Bahara, H., Kranenberg, A., & Tokmetzis, D. (2019, February 8). Hoe YouTube rechtse radicalisering in de hand werkt. *De Volkskrant*. https://www.volkskrant.nl/kijkverder/v/2019/hoe-youtube-rechtse-radicalisering-in-de-hand-werkt/

5. Bell, A. (2007). Designing and testing questionnaires for children. *Journal of Research in Nursing, 12*(5), 461–469. https://doi.org/10.1177/1744987107079616

6. Bell, B. T. (2016). Understanding Adolescents. In L. Little, D. Fitton, B. T. Bell, & N. Toth (Eds.*), Perspectives on HCI research with teenagers* (pp. 11–27). Springer Publishing. https://doi.org/10.1007/978-3-319-33450-9_2

7. Bower, M. (2017). *Design of Technology-enhanced Learning: Integrating Research and Practice*. Emerald Publishing.

8. Bozdag, E., & van den Hoven, J. (2015). Breaking the filter bubble: democracy and design. *Ethics and Information Technology, 17*(4), 249–265. https://doi.org/10.1007/s10676-015-9380-y

9. Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry, 189*(194), 4-7.

10. Bruns, A. (2019). Filter bubble. *Internet Policy Review, 8*(4), 1–14. https://doi.org/10.14763/2019.4.1426

11. Bryant, L. V. (2020). The YouTube Algorithm and the Alt-Right Filter Bubble. *Open Information Science, 4*(1), 85–90. https://doi.org/10.1515/opis-2020-0007

12. Burbach, L., Halbach, P., Ziefle, M., & Calero Valdez, A. (2019*). Bubble Trouble: Strategies Against Filter Bubbles in Online Social Networks*. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11582 LNCS, 441–456. https://doi.org/10.1007/978-3-030-22219-2_33

13. Byrne, Z. S., Dvorak, K. J., Peters, J. M., Ray, I., Howe, A., & Sanchez, D. (2016). From the user's perspective: Perceptions of risk relative to benefit associated with using the Internet. *Computers in Human Behavior, 59*, 456–468. https://doi.org/10.1016/j.chb.2016.02.024

14. Centraal Bureau voor de Statistiek. (2020). Computerkennis en vaardigheid; persoonskenmerken, 2012-2019. https://www.cbs.nl/nl-nl/cijfers/detail/83428NED?dl=3393C

15. Centraal Bureau voor de Statistiek. (2020). Internet jongeren; toegang, gebruik, faciliteiten en online aankopen. StatLine Jeugdmonitor. https://jmopendata.cbs.nl/#/JM/nl/dataset/20255NED/table?dl=3AD63

16. Centraal Bureau voor de Statistiek. (2020). Deelname onderwijs; jongeren 0 tot 25 jaar, regio's. StatLine Jeugdmonitor. https://jmopendata.cbs.nl/#/JM/nl/dataset/71010ned/table?dl=4766B %0A

17. Chitra, U., & Musco, C. (2020). Analyzing the impact of filter bubbles on social network polarization. *WSDM 2020 - Proceedings of the 13th International Conference on Web Search and Data Mining*, 115–123. https://doi.org/10.1145/3336191.3371825

18. Chrysanthou, A., Barlas, P., Kyriakou, K., Kleanthous, S., & Otterbacher, J. (2020). Bursting the bubble: Tool for awareness and research about overpersonalization in information access systems. *International Conference on Intelligent User Interfaces, Proceedings IUI*, 112–113. https://doi.org/10.1145/3379336.3381863

19. Courtois, C., Slechten, L., & Coenen, L. (2018). Challenging Google Search filter bubbles in social and political information: Disconfirming evidence from a digital methods case study. *Telematics and Informatics, 35*(7), 2006–2015. https://doi.org/10.1016/j.tele.2018.07.004

20. Curriculum.nu. (2019). *Leergebied Digitale Geletterdheid*. https://www.curriculum.nu/voorstellen/digitale-geletterdheid/

21. Dennen, V. P., Choi, H., & Word, K. (2020). Social media, teenagers, and the school context: a scoping review of research in education and related fields. *Educational Technology Research and Development, 68*(4), 1635–1658. https://doi.org/10.1007/s11423-020-09796-z

22. DiFranzo, D., Choi, Y. H., Purington, A., Taft, J. G., Whitlock, J., & Bazarova, N. N. (2019). Social Media TestDrive. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Published. https://doi.org/10.1145/3290605.3300533

23. Dunham, R. B. (2006). *Nominal group technique*: A user's guide* [Slides]. SSWM. https://sswm.info/sites/default/files/reference_attachments/DUNHAM%201998%20Nominal%20Group%20Technique%20-%20A%20Users'%20Guide.pdf

24. Eliasson, J., Pargman, T. C., Nouri, J., Spikol, D., & Ramberg, R. (2011). Mobile Devices as Support Rather than Distraction for Mobile Learners. *International Journal of Mobile and Blended Learning, 3*(2), 1–15. https://doi.org/10.4018/jmbl.2011040101

25. Escueta, M., Nickow, A. J., Oreopoulos, P., & Quan, V. (2020). Upgrading Education with Technology : Insights from Experimental Research. *Journal of Economic Literature, 58*(4), 897–996. https://doi.org/10.1257/jel.20191507

26. Fitton, D., & Bell, B. (2014). Working with teenagers within HCI research: Understanding teen-computer interaction. *Proceedings of the 28th International BCS Human Computer Interaction Conference: Sand, Sea and Sky - Holiday HCI, HCI 2014*, 201–206.

27. Fitton, D., Bell, B. T., Little, L., Horton, M., Read, J. C., Rouse, M., & Toth, N. (2016). Working with Teenagers in HCI Research: A Reflection on Techniques Used in the Taking on the Teenagers Project. In L. Little, D. Fitton, B. T. Bell, & N. Toth (Eds.), *Perspectives on HCI Research with Teenagers* (pp. 237–267). Springer Publishing. https://doi.org/10.1007/978-3-319-33450-9_10

28. Fitton, D., Read, J., & Horton, M. (2013). The Challenge of Working with Teens as Participants in Interaction Design. *Proceedings of CHI 2013: Changing Perspectives* (pp. 205–210). ACM. https://doi.org/10.1145/2468356.2468777

29. Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly, 80*(Special issue 1), 298–320. https://doi.org/10.1093/poq/nfw006

30. Fouquaert, T. (2019). The Algorithm Paradox: knowledge of versus attitudes towards curation algorithms [Master thesis]. https://www.scriptieprijs.be/sites/default/files/thesis/2019-10/Thesis_Thibault_Fouquaert_Algorithm_Paradox.pdf

31. Geschke, D., Lorenz, J., & Holtz, P. (2019). The triple-filter bubble: Using agent-based modelling to test a meta-theoretical framework for the emergence of filter bubbles and echo chambers. *British Journal of Social Psychology, 58*(1), 129–149. https://doi.org/10.1111/bjso.12286

32. Gran, A. B., Booth, P., & Bucher, T. (2020). To be or not to be algorithm aware: a question of a new digital divide*? Information Communication and Society*, 1–18. https://doi.org/10.1080/1369118X.2020.1736124

33. Hansen, E. I. K., & Iversen, O. S. (2013). You are the real experts! - Studying teenagers' motivation in Participatory Design. *ACM International Conference Proceeding Series*, 328–331. https://doi.org/10.1145/2485760.2485826

34. Hitlin, P., & Rainie, L. (2019). Facebook Algorithms and Personal Data. *Pew Research Center*, 1–23. http://www.pewinternet.org/2019/01/16/facebook-algorithms-and-personal-data/

35. Horkoff, J., Ersare, J., Kahler, J., Jorundsson, T. D., & Hammouda, I. (2018). Efficiency and effectiveness of requirements elicitation techniques for children. *Proceedings - 2018 IEEE 26th International Requirements Engineering Conference, RE 2018*, 194–204. https://doi.org/10.1109/RE.2018.00028

36. Hsieh, Y. H., Lin, Y. C., & Hou, H. T. (2016). Exploring the role of flow experience, learning performance and potential behavior clusters in elementary students' game-based learning. *Interactive Learning Environments, 24*(1), 178–193. https://doi.org/10.1080/10494820.2013.834827

37. Huisman, C. (2020, September 9). Een tablet of laptop eisen van leerlingen – het mag niet, maar scholen komen ermee weg. *De Volkskrant*. https://www.volkskrant.nl/nieuws-achtergrond/een-tablet-of-laptop-eisen-van-leerlingen-het-mag-niet-maar-scholen-komen-ermee-weg~b7b3b812/

38. Hussein, E., Juneja, P., & Mitra, T. (2020). Measuring Misinformation in Video Search Platforms: An Audit Study on YouTube. *Proceedings of the ACM on Human-Computer Interaction, 4*(CSCW1). https://doi.org/10.1145/3392854

39. Iivari, N., Kinnula, M., Kuure, L., & Keisanen, T. (2020). "Arseing around was Fun!" A- Humor as a Resource in Design and Making. *Conference on Human Factors in Computing Systems - Proceedings*, 1–13. https://doi.org/10.1145/3313831.3376169

40. Iversen, O. S., Dindler, C., & Hansen, E. I. K. (2013). Understanding teenagers' motivation in participatory design. *International Journal of Child-Computer Interaction, 1*(3–4), 82–87. https://doi.org/10.1016/j.ijcci.2014.02.002

41. Iversen, O. S., & Smith, R. C. (2012). Scandinavian participatory design - Dialogic curation with teenagers. *ACM International Conference Proceeding Series*, 106–115. https://doi.org/10.1145/2307096.2307109

42. Jackson, S. A., Martin, A. J., & Eklund, R. C. (2008). Long and short measures of flow: The construct validity of the FSS-2, DFS-2, and new brief counterparts. *Journal of Sport and Exercise Psychology, 30*(5), 561–587. https://doi.org/10.1123/jsep.30.5.561

43. Jin, F., & Divitini, M. (2020). Affinity for Technology and Teenagers' Learning Intentions. *ICER 2020 - Proceedings of the 2020 ACM Conference on International Computing Education Research*, 48–55. https://doi.org/10.1145/3372782.3406269

44. Joyce, A., & Nielsen, J. (2019). *Teenager's UX: Designing for Teens*. Nielsen Norman Group. https://www.nngroup.com/articles/usability-of-websites-for-teenagers/

45. Kinnula, M., & Iivari, N. (2019). Empowered to Make a Change: Guidelines for empowering the young generation in and through digital technology design. *Proceedings of the FabLearn Europe 2019 Conference on ZZZ - FabLearn Europe '19*. Published. https://doi.org/10.1145/3335055.3335071

46. Kirschner, P. A., & De Bruyckere, P. (2017). The myths of the digital native and the multitasker. *Teaching and Teacher Education, 67*, 135–142. https://doi.org/10.1016/j.tate.2017.06.001

47. Kolb, L. (2017). *Learning First, Technology Second: The Educator's Guide to Designing Authentic Lessons.* International Society for Technology in Education.

48. Liao, C., Guimbretière, F., Anderson, R., Linnell, N., Prince, C., & Razmov, V. (2007). PaperCP: Exploring the integration of physical and digital affordances for active learning. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 4663 LNCS(PART 2), 15–28. https://doi.org/10.1007/978-3-540-74800-7_2

49. Little, L., Fitton, D., Bell, B. T., & Toth, N. (2016). *Perspectives on HCI Research with Teenagers*. Springer.

50. Mahrt, M. (2020). Why we find little evidence of digital fragmentation, but should not stop researching it. In V. Gehrau, A. Waldherr, & A. Scholl (Eds.), *Integration durch Kommunikation (in einer digitalen Gesellschaft): Jahrbuch der Deutschen Gesellschaft für Publizistik- und Kommunikationswissenschaft 2019* (pp. 27-35). Münster: Deutsche Gesellschaft für Publizistik- und Kommunikationswissenschaft e.V. https://doi.org/10.21241/ssoar.66412

51. Maqsood, S., Mekhail, C., & Chiasson, S. (2018). A day in the life of Jos: A web-based game to increase children's digital literacy. *IDC 2018 - Proceedings of the 2018 ACM Conference on Interaction Design and Children*, 241–252. https://doi.org/10.1145/3202185.3202753

52. Miglino, O., Di Ferdinando, A., Di Fuccio, R., Rega, A., & Ricci, C. (2014). Bridging digital and physical educational games using RFID/NFC technologies. *Journal of E-Learning and Knowledge Society, 10*(3), 89–106. https://doi.org/10.20368/1971-8829/959

53. Milan, S., & Agosti, C. (2019). *Personalisation algorithms and elections: Breaking free of the filter bubble.* Web publication/site, Internet Policy Review. https://policyreview.info/articles/news/personalisation-algorithms-and-elections-breaking-freefilter-bubble/1385

54. Munson, S. A., Lee, S. Y., & Resnick, P. (2013). Encouraging reading of diverse political viewpoints with a browser widget. *Proceedings of the 7th International Conference on Weblogs and Social Media, ICWSM 2013, Festinger 1957*, 419–428.

55. Mutsaers, P., & Demir, S. (2019). *Preventing radicalization and polarization: A literature review of policies, priorities, programs, and partnerships for youth at risk.* Web publication/site, Platform JEP. https://www.platformjep.nl/binaries/platformjep/documenten/publicaties/2020/04/09/preventing-radicalization-and-polarization/Mutsaers+and+Demir+%282019%29.pdf

56. Nagulendra, S., & Vassileva, J. (2014). Understanding and controlling the filter bubble through interactive visualization: A user study. *HT 2014 - Proceedings of the 25th ACM Conference on Hypertext and Social Media*, 107–115. https://doi.org/10.1145/2631775.2631811

57. Nechushtai, E., & Lewis, S. C. (2019). What kind of news gatekeepers do we want machines to be? Filter bubbles, fragmentation, and the normative dimensions of algorithmic recommendations. *Computers in Human Behavior, 90*(January 2019), 298–307. https://doi.org/10.1016/j.chb.2018.07.043

58. Nederlands Jeugdinstituut. (2018). *Factsheet voor beroepskrachten.* https://www.nji.nl/nl/Download-NJi/Publicatie-NJi/Polarisatie,-radicalisering,-extremisme-en-mediaopvoeding-brochure.pdf

59. NEWCOM. (2020). *Nationale Social Media Onderzoek 2020*. https://www.newcom.nl/socialmedia2020

60. Nieuwelink, H. (2020). Jongeren, media en democratie: Wat adolescenten vinden en onderwijs kan bijdragen. In J. de Ridder, R. Vliegenthart, & J. Zuure (Eds.), *Doen, durven of de waarheid?: Democratie in digitale tijden* (pp. 136-150). Amsterdam University Press.

61. O'Callaghan, D., Greene, D., Conway, M., Carthy, J., & Cunningham, P. (2013*). The Extreme Right Filter Bubble*. Retrieved from http://arxiv.org/abs/1308.6149

62. Overbeek, G., & van Rijn-van Gelderen, L. (2019). Jeugd en media. *Kind En Adolescent, 40*(2), 101–103. https://doi.org/10.1007/s12453-019-00202-x

63. Pariser, E. (2011). *The Filter Bubble: What the Internet is Hiding from You*. New York, NY: The Pinguin Press.

64. Piper, A. M., & Hollan, J. D. (2009). Tabletop displays for small group study: Affordances of paper and digital materials. *Conference on Human Factors in Computing Systems - Proceedings*, 1227–1236. https://doi.org/10.1145/1518701.1518885

65. Platform JEP. (2019). *Jongeren en polarisatie.* https://www.platformjep.nl/documenten/publicaties/2019/07/17/verkenning-jongeren-en-polarisatie

66. Plettenberg, N., Nakayama, J., Belavadi, P., Halbach, P., Burbach, L., Calero Valdez, A., & Ziefle, M. (2020). User behavior and awareness of filter bubbles in social media. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12199 LNCS*, 81–92. https://doi.org/10.1007/978-3-030-49907-5_6

67. Poels, K., De Kort, Y. A. W., & IJsselsteijn, W. A. (2007). *D3.3 : Game Experience Questionnaire: development of a self-report measure to assess the psychological impact of digital games*. Technische Universiteit Eindhoven.

68. Poole, E. S., & Peyton, T. (2013). Interaction design research with adolescents: Methodological challenges and best practices. *ACM International Conference Proceeding Series*, 211–217.

69. Powers, E. (2017). My News Feed is Filtered?: Awareness of news personalization among college students. *Digital Journalism, 5*(10), 1315–1335. https://doi.org/10.1080/21670811.2017.1286943

70. Raby, R. (2010). Public Selves, Inequality, and Interruptions: The Creation of Meaning in Focus Groups with Teens. *International Journal of Qualitative Methods, 9*(1), 1–15. https://doi.org/10.1177/160940691000900101

71. Rader, E., & Gray, R. (2015). Understanding user beliefs about algorithmic curation in the Facebook news feed. *Conference on Human Factors in Computing Systems - Proceedings, 2015-April*, 173–182. https://doi.org/10.1145/2702123.2702174

72. RAN. (2018). *Transforming schools into labs for democracy: A companion to preventing violent radicalisation through education*. https://ec.europa.eu/home-affairs/sites/homeaffairs/files/what-we-do/networks/radicalisation_awareness_network/about-ran/ran-edu/docs/ran_edu_transforming_schools_into_labs_for_democracy_2018_en.pdf

73. Read, J., & Horton, M. (2016). Future Directions for Quality TeenCI Research. In L. Little, D. Fitton, B. T. Bell, & N. Toth (Eds.), *Perspectives on HCI Research with Teenagers* (pp. 269–278). Springer Publishing.

74. Resnick, P., Garrett, R. K., Kriplean, T., Munson, S. A., & Stroud, N. J. (2013). Bursting your (filter) bubble. *CSCW'13: Proceedings of the 2013 conference on Computer supported cooperative work companion,* 95-100. https://doi.org/10.1145/2441955.2441981

75. Roschelle, J., Penuel, W. R., Yarnall, L., & Tatar, D. (2004). Handheld tools that "informate" assessment of student learning in science: A requirements analysis. *Proceedings - 2nd IEEE International Workshop on Wireless and Mobile Technologies in Education*, 149–153. https://doi.org/10.1109/WMTE.2004.1281367

76. Roschelle, J., Shechtman, N., Tatar, D., Hegedus, S., Hopkins, B., Empson, S., Knudsen, J., & Gallagher, L. P. (2010). Integration of technology, curriculum, and professional development for advancing middle school mathematics: Three large-scale studies. *American Educational Research Journal, 47*(4), 833–878. https://doi.org/10.3102/0002831210367426

77. Rose, E. J., Björling, E. A., Kim, A., & Alvarez, N. Y. (2018). Usability testing with teens: Adapting human-centered design and UX methods. *SIGDOC 2018 - 36th ACM International Conference on the Design of Communication,* 1–6. https://doi.org/10.1145/3233756.3233955

78. Seargeant, P., & Tagg, C. (2019). Social media and the future of open debate: A user-oriented approach to Facebook's filter bubble conundrum. *Discourse, Context and Media, 27*, 41–48. https://doi.org/10.1016/j.dcm.2018.03.005

79. Shin, W., & Lwin, M. O. (2017). How does "talking about the Internet with others" affect teenagers' experience of online risks? The role of active mediation by parents, peers, and school teachers. *New Media and Society, 19*(7), 1109–1126. https://doi.org/10.1177/1461444815626612

80. Sindermann, C., Elhai, J. D., Moshagen, M., & Montag, C. (2020). Age, gender, personality, ideological attitudes and individual differences in a person's news spectrum: how many and who might be prone to "filter bubbles" and "echo chambers" online? *Heliyon, 6*(1), e03214. https://doi.org/10.1016/j.heliyon.2020.e03214

81. Spohr, D. (2017). Fake news and ideological polarization: Filter bubbles and selective exposure on social media. *Business Information Review, 34*(3), 150–160. https://doi.org/10.1177/0266382117722446

82. Stroud, N. J. (2008). Media use and political predispositions: Revisiting the concept of selective exposure. *Political Behavior, 30*(3), 341–366. https://doi.org/10.1007/s11109-007-9050-9

83. Stroud, N. J. (2010). Polarization and partisan selective exposure. *Journal of Communication, 60*(3), 556–576. https://doi.org/10.1111/j.1460-2466.2010.01497.x

84. Tatar, D., Roschelle, J., Vahey, P., & Penuel, W. R. (2003). Handhelds Go to School. *IEEE Computer, 36*(9), 30–37.

85. Terrenghi, L., Kirk, D., Sellen, A., & Izadi, S. (2007). Affordances for manipulation of physical versus digital media on interactive surfaces. *Conference on Human Factors in Computing Systems - Proceedings*, 1157–1166. https://doi.org/10.1145/1240624.1240799

86. Thijs, A., Fisser, P., & Der, Van Hoeven, M. (2014). *21E Eeuwse vaardigheden in het curriculum van het funderend onderwijs*. Enschede: SLO. http://downloads.slo.nl/Repository/21e-eeuwse-vaardigheden-in-het-curriculum-van-het-funderend-onderwijs.pdf

87. Van Bergen, D. D., de Ruyter, D. J., & Pels, T. V. M. (2017). "Us Against Them" or "All Humans Are Equal": Intergroup Attitudes and Perceived Parental Socialization of Muslim Immigrant and Native Dutch Youth. *Journal of Adolescent Research, 32*(5), 559–584. https://doi.org/10.1177/0743558416672007

88. Watson, D., & Clark, L. A. (1994). Manual for the positive and negative affect schedule-expanded form. *Unpublished manuscript, University of Iowa, Iowa City*.

89. Wilkinson, C. R. (2016). Involving Teenagers Today in the Design of Tomorrow's Technology. In L. Little, D. Fitton, B. T. Bell, & N. Toth (Eds.), *Perspectives on HCI Research with Teenagers* (pp. 179–206). Springer Publishing.

90. Yap, C. E. L., & Lee, J. J. (2020). “Phone apps know a lot about you!”: Educating early adolescents about informational privacy through a phygital interactive book. *Proceedings of the Interaction Design and Children Conference, IDC 2020*, 49–62. https://doi.org/10.1145/3392063.3394420

91. Yarosh, S., Radu, I., Hunter, S., & Rosenbaum, E. (2011). Examining values: An analysis of nine years of IDC research. *Proceedings of IDC 2011 - 10th International Conference on Interaction Design and Children*, 136–144. https://doi.org/10.1145/1999030.1999046

92. Zuckerman, O., & Gal-Oz, A. (2013). To TUI or not to TUI: Evaluating performance and preference in tangible vs. graphical user interfaces. *International Journal of Human Computer Studies, 71*(7–8), 803–820. https://doi.org/10.1016/j.ijhcs.2013.04.003

93. Zuiderveen Borgesius, F. J., Trilling, D., Möller, J., Bodó, B., De Vreese, C. H., & Helberger, N. (2016). Should we worry about filter bubbles? *Internet Policy Review, 5*(1), 1–16. https://doi.org/10.14763/2016.1.401

Appendix A. Interview protocol for pre-study 1

Introduction

Heel fijn dat je mee wil werken met dit interview. Wij zijn Tim en Anneleen en we studeren Human-Computer Interaction aan de Universiteit Utrecht. Voor onze masterscripties doen we onderzoek naar jongeren en filterbubbels. Hiervoor interviewen we docenten en leerlingen van de middelbare school. Jij bent als docent expert op het gebied van lesgeven en jongeren, daarom kunnen jouw kennis en ervaring ons goed helpen.

We zouden dit gesprek graag willen opnemen. De antwoorden die je geeft zullen anoniem verwerkt worden in onze scripties. Je kan op ieder moment stoppen met dit interview, zonder daarvoor een reden te geven. Daarnaast zullen we altijd je gegevens verwijderen als je daarom vraagt. Geef je toestemming voor het opnemen van dit interview?

Een filterbubbel ontstaat als social media, zoekmachines of andere websites hun inhoud aanpassen op jouw persoonlijke voorkeuren. Daardoor krijg je meer van hetzelfde te zien en worden andere perspectieven, meningen of kritieken weggefilterd. Je kan zelf ook bijdragen aan je filterbubbel, bijvoorbeeld als je altijd dezelfde soort websites bezoekt of alleen bepaalde berichten liket.

Questions

1. Kan je ons eerst wat vertellen over jezelf: welk vak geef je op welk niveau en hoe lang doe je dat al? Welke opleiding heb je gedaan?
2. Kan je ons vertellen wat de rol van technologie is in jouw lessen?
3. Wat is jouw beeld van het online gedrag van leerlingen?
4. Hoe gaan leerlingen om met de informatie die zij online vinden?
5. Welke vorm van lesgeven wekt het meeste enthousiasme op bij leerlingen?
6. Op welke manier zou jij een onderwerp als polarisatie bespreken of behandelen in de klas?
    a. Kan je omschrijven waar je op let? Wat de randvoorwaarden zijn van een gesprek of discussie?
7. Hoe zou technologie je hierbij kunnen ondersteunen?

Voor onze scripties maken we een educatieve app over filterbubbels. Het doel van de app is leerlingen bewust maken van het feit dat ze in een filterbubbel zitten en welke consequenties dit heeft. De filter bubbel is een redelijk abstract concept, maar ligt ook dicht bij de belevingswereld van jongeren.

8. Op wat voor manier zou je zo'n soort concept overbrengen in de les?
    a. Wat werkt er voor jongeren volgens jou beter: een verhalende aanpak of een meer directe aanpak?
9. Hoe zie je voor je dat een app gebruikt kan worden in de les? (bijvoorbeeld: aantal leerlingen, groepjes, gezamenlijk/alleen)
    a. Hoe vaak zou je de applicatie willen gebruiken? (bijvoorbeeld: 1 les, meerdere weken)
10. Op wat voor manier zou jij als docent een rol willen spelen tijdens een les met de applicatie?
11. Heb je nog vragen of opmerkingen? Wil je nog terugkomen op eerdere vragen of antwoorden?

Additional list of topics

- Op wat voor manier wordt er bij jullie op school aandacht besteed aan thema's als burgerschap, digitale geletterdheid, en mediawijsheid?
- Merk je ook iets van bubbels in de klas?
- Wat is de invloed van een docent bij het bespreken van een gevoelig thema?

- Wat zijn vmbo-specifieke eigenschappen, wat daar wel doen, wat zijn juist de valkuilen?
- Wat vind jij van de neutraliteit van de docent?

Additional topics for E1

1. Kan je iets over jezelf vertellen? Wat doe je precies? Hoe lang al?
2. Met wat voor scholen werk je meestal samen, bijvoorbeeld qua niveau?

- Technologie en onderwijs
    - Kansen, uitdagingen, problemen
    - Apps in de klas
- Docenten
    - Waar lopen ze tegenaan?
    - Wat vinden ze leuk?
    - Houding t.o.v. technologie
- Samenwerken in grote projecten
    - Hoe houd je iedereen tevreden?
    - Hoe let je op alle belangen?
- Kinderen
    - Heb je ervaring met ze betrekken tijdens project?
- Algemeen
    - Do's en don'ts, tips & tricks

Appendix B. Insights from pre-study 1

These 40 insights were generated from the interviews with five teachers and one expert, as described in Chapter 4. These interviews generated 397 phrases that we categorized in 14 different groups, from which we distilled the following 40 insights.

What do students do online?

- Students want to keep social media to themselves and don't often like to share it with parents and/or teachers.
- Students are strongly connected to each other online.
- Whether a teacher knows what their students are doing online heavily depends on the teacher.

How do students handle online information?

- People often think that students/teenagers can easily handle information and/or technology. This assumption is wrong.
- Students really do struggle with searching and finding information online, and often just trust everything they find.
- Students do not have a lot of knowledge of current events. This has decreased over the years. Students are not up to date. Even some big news events aren't always properly followed.

Students in general

- The term 'students' is too broad, because there are too many individual differences.
- Vmbo-students rather have 'something to do', are distracted more easily, and will work less for themselves but more for the teacher.
- Students should first recognise the importance of something (for themselves) before they'll want to do something with it.

Teaching in general

- Use a lot of variation, especially at vmbo.
- Competition always works well.
- Lessons should be meaningful for the students, by for example using examples directly from their personal lives.
- It helps for students to work towards a bigger goal; that could be something like developing a product.
- Students become enthusiastic if they notice they are learning something and they can do something they couldn't do before.

Role of technology in class

- If students have access to devices (such as Chromebooks), using technology in class becomes a lot easier.
- Quizzes like Kahoot, Mentimeter and Socrative are often used, either to measure prior knowledge or to gather viewpoints and opinions.
- Technology isn't often used as a means of interaction, but as a way to gather knowledge.

Talking about polarisation

- Activating prior knowledge is important.
- A safe environment in class is fundamental when talking about sensitive topics.
- Focussing on knowledge and formulating opinions based on facts and numbers can help during discussions.
- It may help to gather individual opinions and viewpoints beforehand and then discuss those points. Otherwise, the loud minority might define the entire discussion.

- It needs to be made sure that the discussion does not increase the differences in a class and it does not strengthen people in their own opinion. That is the exact opposite of what should be achieved.

Technology and polarisation

- When discussing polarisation, the teacher can make relevant online sources available for students through technology (for example with QR codes in assignments).
- Adaptive forms of learning, that adapt to the student, can be useful in this context.

Conveying the concept of a filter bubble

- Items such as *Zondag met Lubach* or *The Social Dilemma* (Netflix) are popular among students. However, these items are often only about the extremes, so they need to be 'translated' to something more relatable for students.
- Whether a direct or a narrative approach works best heavily depends on the type of student.
- Students should experience something in different, realistic contexts with appealing examples. That will make it more meaningful for a student ('context-concept approach').

The application in class

- The application should allow students to experience something, and not just convey knowledge.
- Students should be able to share things with each other, which means working in (small) groups is preferable. Working individually is not as effective. Differences between group members can elicit curiosity.
- Differentiating between students is important: some students might be interested, some won't be, some already know about filter bubbles and some not at all. They should all benefit from the application. Letting students make choices themselves can also be helpful.
- In terms of usability, the application should be able to be understood at first glance, for both students and teachers.

Curriculum/place of the application in teaching

- The application should return multiple times, using it once is not going to help at all.
- The application should not be stand-alone; a guide for teachers, a course package or lesson plans are a necessity.
- In general, media/digital literacy should get a structural/more prominent place in education.
- Teachers mention different possibilities regarding where to use the application. It could be used in weeklong projects, in the mentor programme or in courses like civics, philosophy of life, or computer science.

Role of the teacher

- Teachers have a very important role in leading discussions and maintaining a safe atmosphere. Teachers should be able to intervene and should also have the capabilities to do that.
- For teachers, it is sometimes difficult to keep an overview of what all students are doing within an application. Teachers could be supported by providing for example a dashboard or partitioning the application into smaller parts.
- Teachers do not have to stay neutral in discussions and can certainly show their opinion on some sensitive topics, but should make clear that their opinion isn't the only opinion around and that students aren't expected to have the same opinion.
- Teachers mention different roles they expect teachers to have when working with the application: initiator, supervisor or just having the application as a teacher (implying they themselves do not have to do anything). Their role will entirely depend on the format of the application itself.
- Especially on vmbo, the social relationship of the teacher with students is very important.

Appendix C. Interview protocol and topic list pre-study 2

Heel fijn dat je mee wil werken met dit interview. Wij zijn Tim en Anneleen en we studeren Human-Computer Interaction aan de Universiteit Utrecht. Voor onze masterscripties doen we onderzoek naar jongeren en filterbubbels. Hiervoor interviewen we docenten en leerlingen van de middelbare school. Jij bent als docent expert op het gebied van lesgeven en jongeren, daarom kunnen jouw kennis en ervaring ons goed helpen.

We zouden dit gesprek graag willen opnemen. De antwoorden die je geeft zullen anoniem verwerkt worden in onze scripties. Je kan op ieder moment stoppen met dit interview, zonder daarvoor een reden te geven. Daarnaast zullen we altijd je gegevens verwijderen als je daarom vraagt. Geef je toestemming voor het opnemen van dit interview?

Een filterbubbel ontstaat als social media, zoekmachines of andere websites hun inhoud aanpassen op jouw persoonlijke voorkeuren. Daardoor krijg je meer van hetzelfde te zien en worden andere perspectieven, meningen of kritieken weggefilterd. Je kan zelf ook bijdragen aan je filterbubbel, bijvoorbeeld als je altijd dezelfde soort websites bezoekt of alleen bepaalde berichten liket.

---

1. Kan je ons eerst wat vertellen over jezelf: welk vak geef je op welk niveau en hoe lang doe je dat al?
2. Kan je ons vertellen wat de rol van technologie is in jouw lessen?

3. Wat is jouw beeld van het online gedrag van leerlingen?
4. Hoe gaan leerlingen om met de informatie die zij online vinden?
5. Welke vorm van lesgeven wekt het meeste enthousiasme op (bij VMBO leerlingen)?
6. Jij bent ervaringsdeskundige op het gebied van lesgeven en jongeren. Op welke manier zou jij een onderwerp als polarisatie bespreken of behandelen in de klas?
    a. Kan je omschrijven waar je op let? Wat de randvoorwaarden zijn van een gesprek of discussie?

Storyboard
Voor onze scripties maken we een educatieve app over filterbubbels. Het doel van de app is leerlingen bewust maken van het feit dat ze in een filterbubbel zitten en welke consequenties dit heeft. De filter bubbel is een redelijk abstract concept, maar ligt ook dicht bij de belevingswereld van jongeren.

Op basis van de interviews die we met docenten hebben gedaan, hebben we een eerste idee voor de app. Dit hebben we weergegeven in een storyboard, je ziet op verschillende schermen hoe de app in de klas gebruikt wordt. De 2 storyboards die we hebben gemaakt laten beide een mogelijke activiteit/opdracht zien die in de app zit.

We laten je nu graag het storyboard zien. Houdt er rekening mee dat dit nog een concept is en het simpele ontwerpen zijn. We vertellen hoe de opdracht werkt, stop ons vooral als je vragen hebt of denkt: dit gaat niet werken.

Delen storyboard A (Anneleen)
Delen storyboard B (Tim)

7. Wat vind je van dit idee? Zou het werken?
8. Op wat voor manier zou jij als docent een rol willen spelen tijdens een les met de applicatie?

Afsluiting

9. Heb je nog vragen of opmerkingen? Wil je nog terugkomen op eerdere vragen of antwoorden?

Extra topics

- Merk je ook iets van bubbels in de klas?
- Wat is de invloed van een docent bij het bespreken van een gevoelig thema?
- VMBO-specifieke eigenschappen, wat daar doen, wat zijn de valkuilen?
- Wat vind jij van de neutraliteit van de docent?
- Op wat voor manier wordt er bij jullie op school aandacht besteed aan thema's als burgerschap, digitale geletterdheid, en mediawijsheid?
- App gebruiken in de les: in groepjes?
- Plek app in curriculum
- Op wat voor manier zou je zo'n (bubbel) soort concept overbrengen in de les?
- Wat werkt er voor jongeren volgens jou beter: een verhalende aanpak of een meer directe aanpak?

Appendix D. Insights from pre-study 2

Students

- Pygmalion effect: if you expect that mavo students work less independently, they will work less independently.

Lessons

- Setting learning goals is important, and benefits students.
- Surprising elements or a shock-effect are effective in delivering information.

Role of technology

- Teachers have different opinions about whether to provide students with prepared online sources, or let them search the internet themselves.

Discussing polarisation

- Anonymity can be a tool to discuss sensitive topics in class, to gather opinions without people feeling threatened, embarrassed, or ashamed.

App in class

- Both before and after the use of the application, there should be plenty of room for discussion.

Role of the teacher

- Aside from being able to handle discussions about sensitive topics, teachers should also have sufficient knowledge on the topic. This is not always the case right now.

Storyboard 1 - Positive

- Working in groups is an effective way to do this activity, since you get results from each group, and they can discuss the results within their group.
- Using a phone and watching videos is attractive for students.
- The activity being tiered (small groups -> all groups -> in-class discussion) is good.

Storyboard 1 - Notes and observations

- The way an activity is introduced is very important (activate prior knowledge)
- It might be the case that not all students will want to share their YouTube recommendations, because they might feel that what videos they watch is private information.
- The final in-class discussion with the teacher contains the most important step: there, students should realize what a difference in recommended videos could cause.

Storyboard 1 - Suggestions and tips

- Include a clear instruction for the questions students have to answer after watching the different recommending pages. Concrete is always better.
- The negative side of filter bubbles should not be too underlined; students will take it less seriously if its only negatives.
- Include videos that appeal to students.
- The videos should last 3 or 4 minutes.
- Teachers should have the possibility to see the answers (of open questions) before they are presented on the board. Otherwise, they will be scared by the possibility of students who give weird, funny or provocative answers.

Storyboard 2- Positive

- The shock/surprise effect is clearer in this storyboard.

- Teachers feel this is a nice interactive activity to engage students.
- This activity results in students asking follow-up questions, which is good.

Storyboard 2 - Notes and observations

- Students might find it hard to place themselves in someone else's shoes, especially when this is a fictive character. Using a profile might help students to let go of their own interests.
- If the goal is to empower students by learning how algorithms work, it is not important to teach them the technical details.

Storyboard 2 - Suggestions and tips

- One teacher thinks that this activity is better suited for individual work rather than group work, another thinks group work is better.
- The scores, steps, and explanations in between the moments of choice are important. Students should have a clear idea about the impact of their choice.
- The topic of the videos could be political, but could also be 'polarizing' on topics close to them.
- Be careful that students do not get to see extreme content on YouTube.
- Suggestion to add another game modus where the goal is to show videos that are as diverse as possible.

Manual for teachers

- All teachers mention the need for a 'manual' or at least guidelines, tips, or advice for when using the application in class. Both on a knowledge level (what are algorithms, what are filter bubbles) and on a meta level (how to manage discussions).

Appendix E. Design phase: students participants table

Table 15: Participating students in the design phase. They are grouped in the same composition as the focus groups. 'Kopklas' is in between the last year of primary school and the first of secondary school.

Focus Group	ID	Gender	Grade	Located in
Pilot	PL1	M	2	Amsterdam
	PL2	M	2	Amsterdam
	PL3	M	2	Amsterdam
	PL4	M	2	Amsterdam
FG1	L1	F	3	Utrecht
	L2	F	3	Utrecht
	L3	F	3	Utrecht
FG2	L4	F	1	Utrecht
	L5	M	1	Utrecht
	L6	M	1	Utrecht
	L7	F	1	Utrecht
FG3	L8	M	Kopklas	Amsterdam
	L9	M	Kopklas	Amsterdam
	L10	M	Kopklas	Amsterdam
	L11	F	Kopklas	Amsterdam
	L12	F	Kopklas	Amsterdam
FG4	L13	F	1	Utrecht
	L14	F	1	Utrecht
	L15	F	1	Utrecht
FG5	L16	M	Kopklas	Amsterdam
	L17	M	Kopklas	Amsterdam
	L18	F	Kopklas	Amsterdam
FG6	L19	M	Kopklas	Amsterdam
	L20	F	Kopklas	Amsterdam

Appendix F. Design phase: teacher interview protocol

For this protocol, it is important to realize that not every teacher was shown every activity and/or prototype. All questions are listed below.

We zouden dit gesprek graag willen opnemen. De antwoorden die je geeft zullen anoniem verwerkt worden in onze scripties. Je kan op ieder moment stoppen met dit interview, zonder daarvoor een reden te geven. Daarnaast zullen we altijd je gegevens verwijderen als je daarom vraagt. Geef je toestemming voor het opnemen van dit interview?

Wij zijn Tim en Anneleen en we studeren Human-Computer Interaction aan de Universiteit Utrecht. Binnen het project zijn wij verantwoordelijk voor het ontwerpen van een educatieve app over filterbubbels. Het doel van de app is leerlingen bewust maken van het feit dat ze in een filterbubbel zitten en welke consequenties dit heeft.

We hebben al verschillende docenten van middelbare scholen geïnterviewd en op basis van deze gesprekken hebben we een eerste prototype van de applicatie gemaakt. In het prototype zie je hoe de app in de klas gebruikt wordt. Er zijn op dit moment twee mogelijke activiteiten/opdrachten uitgewerkt die in de app kunnen komen.

We laten je nu graag het prototype zien. Houdt er rekening mee dat dit nog een concept is en het simpele ontwerpen zijn. We vertellen hoe de opdracht werkt, stop ons vooral als je vragen hebt of denkt: dit gaat niet werken.

Een filterbubbel ontstaat als social media, zoekmachines of andere websites hun inhoud aanpassen op jouw persoonlijke voorkeuren. Daardoor krijg je meer van hetzelfde te zien en worden andere perspectieven, meningen of kritieken weggefilterd. Je kan zelf ook bijdragen aan je filterbubbel, bijvoorbeeld als je altijd dezelfde soort websites bezoekt of alleen bepaalde berichten liket.

Introduction

1. Kan je ons eerst wat vertellen over jezelf: welk vak geef je op welk niveau en hoe lang doe je dat al?
2. Kan je ons vertellen wat de rol van technologie is in jouw lessen?
3. Op wat voor manier wordt er bij jullie op school aandacht besteed aan thema's als burgerschap, digitale geletterdheid, en mediawijsheid?

Prototype

4. Wat vind je van dit idee? Zou het werken?
5. Denk je dat leerlingen dit leuk vinden om in de klas te doen?
6. Wat vind je van het niveau van de app?
7. Prototype 1 specifiek:
    a. Wat vind je ervan dat leerlingen hun eigen YouTube account gebruiken?
    b. Hoe voorkom je dat leerlingen filmpjes gaan kijken op YouTube, terwijl dat niet bij de opdracht hoort?
    c. Welke vragen zouden we op het einde moeten stellen?
    d. Wat vind je ervan dat de antwoorden van leerlingen op het bord terecht komen?
8. Prototype 2 specifiek:
    a. Wat vind je van de profielschets vooraf?
    b. Wat vind je van de tekstjes tussendoor?

c. Wat voor type video's moeten gebruikt worden? Polariserend of niet? Iets waar ze neutraal in staan of juist waar ze al een mening over hebben?
d. Hoeveel rondes zouden er moeten zijn?
e. Wat vind je van de live meters op het bord? (en competitie)

9. Prototype 3 specifiek:
    a. Wat vind je van het kiezen tussen de 4 profielen?
        i. Denk je dat leerlingen een beetje verschillend zullen kiezen?
    b. Wat vind je van het onderwerp familievlogs? Meer ideeën voor onderwerpen?
    c. Hoe lang is dit leuk (hoe veel keuzes)?
    d. Denk je dat leerlingen de tekstjes zullen lezen?
    e. Wat denken jullie dat beter is, praten tegen andere leerling of tegen 'computer'?
    f. Wat zou jouw rol als docent zijn tijdens het spelen van deze activiteit? Hoe kunnen wij docenten ondersteunen bij deze activiteit?
10. Prototype platform specifiek:
    a. Wat vinden jullie van dit idee?
    b. Denk je dat leerlingen dit leuk vinden?
    c. Zijn er onderdelen waar jullie direct enthousiast van worden?
    d. Als je dit in de klas zou gebruiken, waar zie je knelpunten of moeilijkheden? Specifiek voor vmbo-leerlingen?
    e. Welke onderdelen hebben extra aandacht nodig? Waar moeten we vooral op letten?
    f. Zou het platform bouwen beter werken in groepjes of individueel? Wat zouden leerlingen zelf leuker vinden?
    g. Hebben jullie ideeën voor het afronden van dit project? Presentatie? Poster maken

Les eromheen

11. Wat vind je van het werken in groepjes met de app?
12. Hoe vaak zou je de app willen gebruiken in de les? Hoeveel lessen?
13. Hoe zou je een les voor je zien met de app?
14. Op wat voor manier zou jij als docent een rol willen spelen tijdens een les met de app?

Afsluiting

15. Heb je nog vragen of opmerkingen? Wil je nog terugkomen op eerdere vragen of antwoorden?

Extra vragen bij tijd over

- Merk je ook iets van bubbels in de klas?
- Wat is de invloed van een docent bij het bespreken van een gevoelig thema?
- Wat zijn VMBO-specifieke eigenschappen, wat werkt, wat zijn de valkuilen?
- Waar kan de app geplaatst worden in het curriculum?
- Op wat voor manier zou je zo'n (bubbel) soort concept overbrengen in de les?

Appendix G. Design phase: students protocol

A. Activity 1 (comparing videos): used in Pilot

- [Op laptop] Wat ik nu ga laten zien, staat normaalgesproken op het bord.
- Uitleg activiteit bij de verschillende schermen.
- Pak nu je telefoon en ga naar de app. Jullie zijn samen een groepje. Normaal is de hele klas verdeeld in groepjes. Je moet tijdens deze opdracht met elkaar overleggen, ik luister gewoon mee. Als je een vraag voor mij hebt, of iets niet snapt mag je dat altijd zeggen.

Vragen

1. Wat vind je ervan dat je je eigen YouTube account gebruikt?
2. Wat vond je van de onderwerpen van de 4 video's waaruit je kon kiezen?
3. Wat vond je van de vragen op het einde?
    a. te moeilijk/makkelijk
    b. veel / weinig
    c. saai / leuk
4. Wat vind je ervan dat de antwoorden van je groepje op het bord komen?
5. Wat vond je ervan om deze opdracht in een groepje te doen?

B. Activity 2 (algorithm activity) Version 1: used in Pilot, FG1

- [Op laptop] Wat ik nu ga laten zien, staat normaalgesproken op het bord.
- Uitleg activiteit bij de verschillende schermen.
- Leerlingen keuzes laten maken over video's

Vragen

1. Is de uitleg van de activiteit duidelijk?
    a. Wil je dit zelf lezen of moet de docent het uitleggen?
2. Wat vind je van de profielschets vooraf?
    a. Wil je meer informatie? Zo ja, wat voor informatie?
    b. Kon je je goed inleven in Robin?
3. Wat vind je van het tekstje voordat je het resultaat te zien krijgt?
    a. Vind je het fijn om van tevoren al iets te weten? Of zie je liever direct het resultaat?
4. Wat vind je van het tekstje bij het resultaat?
    a. Wat vind je van de interesse-score?
    b. Is het duidelijk hoe deze werkt?
5. Wat vond je van de onderwerpen van de video's?
6. Vond je het moeilijk om te kiezen tussen de 2 video's?
    a. Zou je meer video's willen om uit te kiezen?
    b. Is de titel genoeg informatie om de keuze te maken? Wil je meer zien (zoals, likes, thumbnail, kanaal, reacties)
7. Hoeveel rondes zouden er moeten zijn?
    a. Hoe lang zou dit doen leuk blijven?
8. Wat vind je van de live meters op het bord, de competitie met anderen?

Vragen algemeen app

9. Zou je dit leuk vinden om in de klas te doen?
10. Wat zou je veranderen en hoe?
11. Wat zou de app nog leuker maken?
12. Hoe kunnen we de app uitdagender maken?
13. Vind je in groepjes werken met de app leuk?

C. Activity 2 (algorithm activity) Version 2: used in FG2, FG3, FG4

- [Op beamer] Spelen app met de leerlingen samen
- Op papier/bord alvast profielschets invullen in de twee categorieën (leuk/niet leuk)
- Leerlingen keuzes laten maken over video's
- Na iedere stap, op het bord (of groot papier) bijhouden met denkstappen
- Vragen kunnen gaandeweg gesteld worden, of na de app

Vragen tijdens:

- Waarom kiezen jullie voor deze video?
- Waarom kies je niet voor de andere video's?
- Ken je deze video's?
- Kijk je deze video's zelf ook?
- Vind je de video's die YouTube aan jou aanraad leuk?
- Bij welke categorie zou deze video horen?

Doorvragen

- Waarom?
- Wat vind jij daarvan?
- Kan je dat uitleggen?
- Hoe gaat het nu in de klas?
- Denk je dat andere leerlingen daar hetzelfde over denken?
- Kan je hier een voorbeeld van geven?
- Maak je dat zelf ook mee?

Vragen na spelen (keywords marked in bold text)

1. Wil je de **uitleg** van deze activiteit zelf lezen/bekijken in een video of wil je liever dat de docent het uitlegt?
2. Had je genoeg informatie over Robin na het lezen van dit **profiel**?
    a. Wil je meer informatie? Zo ja, wat voor informatie?
    b. Zouden we het profiel ook weg kunnen laten?
3. Vond je het **moeilijk om te kiezen** tussen de 4 video's?
    a. Hoe zou je dit moeilijker maken?
4. We hebben nu 4 **rondes** gespeeld. Zou je nog meer rondes willen spelen? Hoeveel?
5. Wil je dit liever samen spelen in een **groepje** (zoals nu net), of alleen op je eigen laptop of telefoon?
6. Wij zijn nu een groepje met z'n allen. En wij hebben dit spel gespeeld. Als we dit met de hele klas spelen, zijn er natuurlijk meer groepjes. Stel je voor dat iedereen een **score** krijgt aan het einde van het spel. Dus wij krijgen best een hoge score, want we hebben heel veel video's gekozen waar Robin blij van werd. Ieder groepje heeft zo z'n eigen score.
    a. Zou je nog meer je best doen als er scores van verschillende groepjes op het bord staan?
    b. Zou je nog meer je best doen als je kan winnen?
7. Zou je dit **leuk** vinden om in de klas te doen?
8. Hebben jullie nog ideeën hoe we dit **spel** nog **leuker** kunnen **maken**? Jullie spelen misschien zelf ook wel eens spelletjes, of misschien heb je in de les ook wel eens een leuke app of website gebruikt? Is er iets waarvan je denkt, het zou echt cool zijn als dit ook in het spel zat?

D. Activity 3 (filter bubble activity): used in FG5, FG6

- Leerlingen kiezen of ze spelen met Jayden/Sophie
- Lopen zelfstandig het verhaal door tot het einde

Vragen

- Vond je dit leuk om te doen?
- Was het duidelijk? Begreep je het?
- Vond je het moeilijk?
- Had je het gevoel dat je keuzes het verhaal bepaalden?
- Heb je alle tekstjes gelezen? Was het teveel tekst? Denk je dat klasgenoten dit zouden lezen?
- Zou je meer keuzes willen? Of minder? Hoe lang is dit leuk om te doen?

- Herken je dit verhaal zelf ook?
- Heb je ook wel eens meegemaakt dat je haatcomments zag onder een video of dat je in een bubbel kwam? Meer van hetzelfde zag?

<u>Eindgesprek</u>

Stelling: Familievlogs zijn slecht voor kinderen

- Jij was Jayden/Sophie. Wat zouden die er van vinden?
- Denk je dat Jayden/Sophie hetzelfde vindt?
- Waarom komt dit?
- Herken je dit/is dit in het echt ook wel eens gebeurd?
- Vind je het erg dat zoiets kan gebeuren?

<u>Prototype</u>

- Zou je zelf willen typen of is zo een keuze aanklikken fijner?
- Hebben jullie een idee voor andere profielen invulling?
- Is dit een leuk onderwerp/Welke andere onderwerpen zou je willen zien?
- Hebben jullie ideeën hoe we dit leuker/spannender/moeilijker maken?

Appendix H. Activity 2 videos

Below, all videos that were used for Activity 2 in all iterations are listed. Bold titles indicate the 'correct' option. Both video title and channel name are shown.

H.1 First iteration

Robin's profile

- 13 years old
- Plays football
- Plays Minecraft
- Listens to music of Snelle

First video: MINECRAFT Let's Play Aflevering #11

1. Round 1
    a. NEDERLANDSE Make-up Tutorial! | Nikkie Tutorials
    b. **BUILDING a MANSION!! Let's Play Minecraft #52**
2. Round 2
    a. Vaccinatiepaspoort | Zondag met Lubach (S14)
    b. **Playing Fortnite with you guys | PewDiePie**

H.2 Second iteration

Robin's profile

- 13 years old
- Plays football
- Plays Minecraft
- Listens to music of Snelle
- Subscribed to Milan Knol, Supergaande en Touzani

First video: MINECRAFT Let's Play Aflevering #11

1. Round 1
    a. DE BANKZITTERS BESTELLEN ALLES VAN DE SNACKBAR!| Bankzitters
    b. NEDERLANDSE Make-up Tutorial! | Nikkie Tutorials
    c. **Mijn EERSTE DAG in een Minecraft Roleplay server! (Netherlands SMP) #1 | Milan Knol**
    d. 10 Duurste en Zeldzaamste Auto's ter Wereld! | detoptien
2. Round 2
    a. RANDOM *ZELDZAME* SKIN CHALLENGE In Fortnite!| Knijn
    b. **MALINO DE HOND IN MINECRAFT BUILDOFF! #2 | Milan Knol**
    c. Mijn grootste blunder als Imposter in Among Us | GameMeneer
    d. Playing GTA 5 Without BREAKING LAWS For 24 Hours! | Kwebbelkop
3. Round 3
    a. **24 UUR OVERLEVEN IN DE DIERENTUIN (SLAPEN IN HET OLIFANTEN VERBLIJF!)| Milan Knol**
    b. Achtervolging motorscooter door politie Utrecht Noord. | Politievlogger Jan-Willem
    c. The Weeknd - Save Your Tears (Official Music Video) | The Weeknd
    d. Stealing SIREN HEAD CARS In GTA 5 RP! | Kwebbelkop
4. Round 4
    a. IK ZETTE DURE SPULLEN GRATIS OP MARKTPLAATS, EN DIT GEBEURDE ER... | Kalvijn
    b. EEN LEUK CADEAU VAN JADE & HANWE ZIJN HUIS IS AF!😍🏠| Gio
    c. Gaby Blaaser op de Vlucht - Jachtseizoen'20 #6 | StukTV
    d. **Milan Knol op de Vlucht - Jachtseizoen'20 #7 | StukTV**

H.3 Third iteration

Robin's profile

- 13 years old

- Plays football
- Plays Minecraft
- Listens to music of Snelle
- Subscribed to Milan Knol, Supergaande en Touzani
-

First video: Minecraft Challenges #14 - HOE IS DIT MOGELIJK?! | Milan Knol

1. Round 1
    a. DE BANKZITTERS BESTELLEN ALLES VAN DE SNACKBAR!| Bankzitters
    b. NEDERLANDSE Make-up Tutorial! | Nikkie Tutorials
    c. **Mijn EERSTE DAG in een Minecraft Roleplay server! (Netherlands SMP) #1 | Milan Knol**
    d. 10 Duurste en Zeldzaamste Auto's ter Wereld! | detoptien
2. Round 2
    a. RANDOM *ZELDZAME* SKIN CHALLENGE In Fortnite!| Knijn
    b. **MALINO DE HOND IN MINECRAFT BUILDOFF! #2 | Milan Knol**
    c. Mijn grootste blunder als Imposter in Among Us | GameMeneer
    d. Playing GTA 5 Without BREAKING LAWS For 24 Hours! | Kwebbelkop
3. Round 3
    a. **24 UUR OVERLEVEN IN DE DIERENTUIN (SLAPEN IN HET OLIFANTEN VERBLIJF!)| Milan Knol**
    b. Achtervolging motorscooter door politie Utrecht Noord. | Politievlogger Jan-Willem
    c. The Weeknd - Save Your Tears (Official Music Video) | The Weeknd
    d. Stealing SIREN HEAD CARS In GTA 5 RP! | Kwebbelkop
4. Round 4
    a. IK ZETTE DURE SPULLEN GRATIS OP MARKTPLAATS, EN DIT GEBEURDE ER... | Kalvijn
    b. EEN LEUK CADEAU VAN JADE & HANWE ZIJN HUIS IS AF!😍🏠| Gio
    c. Bram Krikke op de Vlucht - Jachtseizoen'19 #1 | StukTV
    d. **Milan Knol op de Vlucht - Jachtseizoen'20 #7 | StukTV**

H.4 Fourth iteration

Robin's profile

- 13 years old
- Subscribed to Milan Knol, Supergaande en Touzani

1. Round 1
    a. DE BANKZITTERS BESTELLEN ALLES VAN DE SNACKBAR!| Bankzitters
    b. NEDERLANDSE Make-up Tutorial! | Nikkie Tutorials
    c. **Mijn EERSTE DAG in een Minecraft Roleplay server! (Netherlands SMP) #1 | Milan Knol**
    d. 10 Duurste en Zeldzaamste Auto's ter Wereld! | detoptien
2. Round 2
    a. RANDOM *ZELDZAME* SKIN CHALLENGE In Fortnite!| Knijn
    b. **FIFA BATTLE CHAMPIONS LEAGUE PSG-BARCE | TOUZANI TV**
    c. Mijn grootste blunder als Imposter in Among Us | GameMeneer
    d. Playing GTA 5 Without BREAKING LAWS For 24 Hours! | Kwebbelkop
3. Round 3
    a. **10 IRRITATIES TIJDENS VOETBAL! | Dylanhaegens**
    b. Achtervolging motorscooter door politie Utrecht Noord. | Politievlogger Jan-Willem
    c. The Weeknd - Save Your Tears (Official Music Video) | The Weeknd
    d. Stealing SIREN HEAD CARS In GTA 5 RP! | Kwebbelkop
4. Round 4
    a. IK ZETTE DURE SPULLEN GRATIS OP MARKTPLAATS, EN DIT GEBEURDE ER... | Kalvijn
    b. EEN LEUK CADEAU VAN JADE & HANWE ZIJN HUIS IS AF!😍🏠| Gio
    c. **Touzani op de Vlucht - Jachtseizoen'19 #6 | StukTV**
    d. **Milan Knol op de Vlucht - Jachtseizoen'20 #7 | StukTV**

Appendix I. Activity 3 storylines

The storylines for Jayden and Sophie are reproduced below in text.

T = text, C = choice, Cx.A/Cx.B = choice options, E = end. The letter and number between parentheses after each part denotes where the player will go next after that part.

I.1 Jayden

T1 Jayden hoort van een vriend over de familie Baaij, ze maken video's over hun leven. Jayden is benieuwd. (C1)

C1 Waar zoekt hij meer informatie?
- C1.A YouTube (T3)
- C1.B Google (T2)

T2 Hij leest dat de Familie Baaij een YouTube kanaal heeft en dagelijks vlogt. (C2)

C2 Wat doet hij nu?
- C2.A Kanaal opzoeken op YouTube (T3)
- C2.B Verder lezen op Google (T4)

T3 Jayden vindt het kanaal van de familie Baaij op YouTube. Hij kijkt wat rond. (C3)

T4 Na wat artikelen over de Familie Baaij en foto's van prijzen die ze hebben gewonnen, komt Jayden uit bij het YouTube kanaal van de Baaij. (C3)

C3 Welke video kijkt hij?
- C3.A Video over de eerste verjaardag van kind (T5)
- C3.B Video over de geboorte van kind (T5)

T5 Jayden leest in een comment dat het helemaal niet goed is voor een kind om zoveel in beeld te zijn. Hij wil weten of dit echt zo is! (C4)

C4 Jayden gebruikt Google en zoekt naar 'kind familievlogs slecht'. Waar klikt Jayden op?
- C4.A Krantenartikel: ‘Kinderen in vlogs: trauma voor het leven’ (T6)
- C4.B Video: Baaij gebruikt kinderen voor geld (T7)

T6 Jayden is in shock door dit artikel. Kinderen in vlogs hebben hier soms nog jaren last van. Als je de hele tijd gefilmd wordt heeft dat veel invloed op je leven. (T8)

T7 Jayden is in shock door deze video. De kinderen van de familie Baaij worden gebruikt om reclame te maken. Dat is toch niet goed voor een kind? (T8)

T8 Jayden is boos. Hij vindt het niet oké dat de kinderen van de familie Baaij zo worden behandeld. Hij besluit er iets van te zeggen. (C5)

C5 Hij klikt een video aan op YouTube. Wat doet hij?
- C5.A Dislike geven (T9)
- C5.B Haatcomment achterlaten (E)

T9 Jayden ziet dat de comments alleen maar positief zijn. Hij wil die mensen vertellen hoe slecht de Baaij vlogs zijn voor de kinderen. (C6)

C6 Wat zet Jayden in zijn comment?
- C6.A 'Denk aan die kinderen! Dit is geen leven zo! STOP HET NU!' (E)
- C6.B 'BELACHELIJK! Dit moet verboden worden!! Weg met de Baaij!' (E)

E Jayden vindt het schandalig en laat haatcomments achter. Hij doet dat ook bij andere video's. Dit moet stoppen!

I.2 Sophie

T1 Sophie scrollt door de YouTube trending pagina en komt een video tegen van de familie Baaij. De thumbnail ziet er aantrekkelijk uit. (C1)

C1 Wat doet Sophie?
- C1.A Video bekijken (T2)
- C1.B Naar het kanaal van de familie Baaij (T3)

T2 Sophie moet erg lachen om wat er allemaal in de video gebeurt. De vader maakt leuke grapjes en de kinderen doen leuke dingen. (C2)

C2 Wat doet Sophie nu?
- C2.A Nog een video aanklikken (T4)
- C2.B Naar het kanaal van de familie Baaij (T3)

T3 Sophie vindt het kanaal van de familie Baaij op YouTube. Ze kijkt wat rond en ziet ook links naar hun andere social media. (C3)

T4 Sophie kijkt nog een video en vindt deze net zo leuk. Aan het einde van de video vraagt de moeder om ze te volgen op social media. Sophie lijkt dit wel leuk. (C3)

C3 Naar welke social media gaat Sophie?
- C3.A Instagram (T5)
- C3.B TikTok (T5)

T5 Sophie scrollt door een paar berichten heen, en vindt het zo leuk dat ze de tijd bijna vergeet. (C4)

C4 Wat doet Sophie?
- C4.A De familie Baaij volgen op social media (T6)
- C4.B Abonneren op de familie Baaij op YouTube (T7)

T6 Sophie kijkt elke dag wel een filmpje dat voorbij komt op social media. Één filmpje vindt ze zo leuk dat ze een comment wil achterlaten. (C5)

T7 Sophie kijkt elk filmpje dat de familie Baaij plaatst op YouTube. Ze ziet dat de dochter een shirt aan heeft dat zij zelf ook heeft en wil daar een comment over achterlaten. (C5)

C5 In de comments staan veel haatreacties en mensen die zeggen dat het zielig is voor kinderen dat ze altijd gefilmd worden. Sophie schrikt hier van. Wat doet ze?
- C5.A Een positieve comment achterlaten (T8)
- C5.B Reageren op een van de haatcomments (C7)

T8 Iemand reageert op haar comment dat ze niet dit soort kanalen moet aanmoedigen en dat ze normaal moet doen. (C6)

C6 Sophie wordt hier boos van en wil reageren. Wat typt ze?
- C6.A ‘Hoezo doe je zo moeilijk, als ik dit leuk vind om te kijken mag dat toch gewoon!’ (E)
- C6.B ‘Bemoei je lekker met je eigen zaken, ik kijk wat ik wil’ (E)

C7 Wat typt Sophie?
- C7.A ‘Die ouders weten toch wel beter dan jij wat hun kinderen leuk vinden!’ (E)
- C7.B ‘Als je dit niet leuk vindt waarom ben je dan hier!’ (E)

E Sophie vindt het niet leuk dat mensen de familievlogs zo aanvallen en reageert fel op de haatcomments in de reacties.

Appendix J. Evaluation questionnaires

J.1 Game Experience Questionnaire (GEQ)

The Game Experience Questionnaire was developed by Poels, De Kort & IJsselstijn (2007). The items in *italics* were adapted either because of mistranslations or to better suit the language level of 12-13 year old VMBO students.

1. Ik was geboeid door het verhaal van het spel
2. Ik voelde me succesvol
3. Ik voelde me verveeld
4. Ik vond het indrukwekkend
5. Ik vergat alles om me heen
6. Ik was gefrustreerd
7. Ik vond het saai
8. Ik was prikkelbaar *(ik was snel geïrriteerd)*
9. *Ik was er goed in*
10. *Ik ging helemaal op in de game*
11. Ik voelde me tevreden
12. Ik voelde me uitgedaagd
13. *Ik had het gevoel dat ik aan het leren was*
14. *Ik voelde me goed*

Answer options:

1. Helemaal niet
2. Een beetje
3. Gemiddeld
4. Best wel
5. Heel erg

J.2 Dispositional Flow Scale (DFS-2)

This short form of the Dispositional Flow Scale was extracted from Jackson, Martin, & Eklund (2008). I translated the items into Dutch, staying as close as possible to the original items, but making sure they were grammatically sound in Dutch.

1. Ik heb het gevoel dat ik genoeg kan voor deze activiteit
2. Ik doe dingen spontaan en automatisch, zonder er over na te hoeven denken
3. Ik weet goed wat ik wil doen
4. Ik weet ik hoe goed ik het aan het doen ben
5. Ik ben compleet geconcentreerd op de activiteit
6. Ik heb de complete controle over wat ik aan het doen ben
7. De tijd lijkt sneller of langzamer te gaan dan normaal
8. De ervaring geeft me een heel tevreden gevoel

Answer options:

1. Nooit (never)
2. Bijna nooit (almost never)
3. Soms (sometimes)
4. Vaak (often)
5. Heel vaak (very often)

J.3 Filter bubble knowledge questions

1. Weet je wat een filterbubbel is?
    a. Ja, en ik kan ook uitleggen wat het (ongeveer) is
    b. Ja, maar ik vind het lastig om het precies uit te leggen
    c. Nee, maar ik heb er wel ooit van gehoord
    d. Nee en ik heb er ook nog nooit van gehoord

2. Kan je in je eigen woorden uitleggen wat een filterbubbel is?
   Als je niet weet wat het is, schrijf je op: ik weet het niet

   (open)

3. Ik zit zelf in een filterbubbel.
   (Helemaal mee oneens -> Helemaal mee eens)

4. Wat mensen online doen of zien kan hun mening veranderen.
   (Helemaal mee oneens -> Helemaal mee eens)

Appendix K. Evaluation protocol and topic list for teachers

The goal of these interviews was to get an expert opinion on whether a physical (paper) artefact included with the application would improve learning or not, both in general and with this specific activity. Additionally, these experts can share their experiences with paper artefacts in class and how students work with these.

Intro
Kort vertellen: ik wil het gaan hebben over het gebruik van papieren hand-outs of werkbladen in de klas.

Questions

1. Maak je wel eens gebruik van papieren hand-outs in je lessen? (werkbladen?)
    a. Wat zijn jouw ervaringen daarmee?
    b. Wat vinden de leerlingen daarvan?
    c. Gebruiken ze ook wel eens papieren hand-outs terwijl ze online iets moeten doen?
2. Wat vind je van het gebruiken van papieren hand-outs?

[Activiteit laten zien, uitleggen hoe daar fysiek vs. digitaal naar voren komt. Ook papieren hand-outs laten zien.]

3. Wat zou voor jou als docent het handigst/fijnst/beste zijn? Waarom?
4. Waar denk je dat leerlingen het meest van zouden leren? Waarom?
5. Het invullen van de bubbel is iets waar de leerlingen zelf aan moeten denken. Gaat dat makkelijker als die bubbel fysiek is of digitaal?
6. Als jij als docent na de activiteit het met de leerlingen gaat hebben over hun opgebouwde bubbels, heb je dan liever dat die bubbels fysiek zijn of digitaal?

Voorbeelden uit praktijk/literatuur om docenten te prikkelen
- Sommige leerlingen die we spraken gaven aan dat papier onhandig is, dat leerlingen dan gaan ruziën om papier, dat het lang duurt om dat uit te delen. Herken jij dat?
- Het zou kunnen dat als leerlingen iets fysieks moeten gebruiken bij iets digitaals, dat dat switchen de hele tijd er voor zorgt dat ze minder geconcentreerd bezig zijn. Hoe denk jij daar over?