The *SpaceStation App*: Design and Evaluation of an AR Application for Educational Television

Simon von der Au
Leon Giering
simonanian.au@campus.lmu.de
leon.giering@campus.lmu.de
Institut für Rundfunktechnik (IRT)
Munich, Germany

Christina Schneegass
christina.schneegass@ifi.lmu.de
LMU Munich
Munich, Germany

Markus Ludwig
markus.ludwig@irt.de
Institut für Rundfunktechnik (IRT)
Munich, Germany

Figure 1: We designed the augmented reality *SpaceStation* app shown in this picture. It provides a hands-on experience and interaction possibilities to viewers of educational television, in this case, a learning video on the International Space Station.

ABSTRACT

Due to the rising popularity of streaming services, television networks are experiencing pressure to keep the attention of the younger audience. Especially in the field of Edutainment, platforms like YouTube or TED are serious competitors and require broadcasters to come up with novel ideas to engage viewers in their program. In this work, we present the augmented reality (AR) *SpaceStation* application, designed to supplement the viewing of educational videos about the ISS. We evaluated users' experience during the interaction with the app in a within-subject user study ($N = 31$) and assessed their workload. During the interaction with the *SpaceStation App*, participants experienced a higher workload compared to a video-only condition; nonetheless, they considered AR a valuable and enjoyable addition. This paper concludes with a discussion from the perspectives of viewers, content creators, and hosts, and states initial ideas on how to design television programs with AR content, without creating information overload.

CCS CONCEPTS

• **Applied computing** → Interactive learning environments; • **Human-centered computing** → **Mixed / augmented reality**.

KEYWORDS

Augmented Reality, Educational Television, Edutainment, Learning

ACM Reference Format:
Simon von der Au, Leon Giering, Christina Schneegass, and Markus Ludwig. 2020. The *SpaceStation App*: Design and Evaluation of an AR Application for Educational Television. In *ACM International Conference on Interactive Media Experiences (IMX '20), June 17–19, 2020, Cornella, Barcelona, Spain.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3391614.3393655

IMX '20, June 17–19, 2020, Cornella, Barcelona, Spain

ACM ISBN 978-1-4503-7976-2/20/06…$15.00
https://doi.org/10.1145/3391614.3393655

1 INTRODUCTION

With the rise of the internet over the last three decades, people's media consumption changed. In Germany, especially younger generations frequently make use of online services such as Netflix, Disney Plus, and YouTube [18], putting long-established television networks under pressure to keep their audience. Not only are streaming services a competition when it comes to entertainment (i.e., movies and series), but also in terms of educational formats (i.e., documentaries, science shows). This so-called *Edutainment* (educational entertainment) content, which is designed to be entertaining, engaging, and supportive of long-term learning [33], is still a common format shown in German television. However, online platforms are steadily growing and attracting a growing number of viewers. For example, the platform TED, which provides videos of short and powerful talks about *T*echnology, *E*ntertainment, and *D*esign, reached 1.5 million views per day in 2012, after just six years of existence [3]. Furthermore, YouTube's educational and learning videos are ranking in the "Top 15" of 2019's most popular video types [17], showing the population's interest in such formats. Being aware of the growing competition in this branch, German broadcasters created edutainment content for other platforms such as YouTube, Instagram, Snapchat, and TikTok (cf. [8, 23]). In the future, they aim to bring these formats together with the traditional program, looking into the direction of: How can we integrate novel technology into the running television program to make it more engaging and interactive?

One possible way is to extend the available content by more interactive technology, such as Augmented Reality (AR) applications. AR can be designed to supplement an existing program and has shown its potential of being both entertaining (e.g., in prominent mobile games like Pokemon Go[1]) and educational. AR has previously been applied in classroom and wider school environments [1] and, besides others, has shown to have a positive impact on the motivation of learners as well as its potential to enrich the learning experience [2, 25]. For the use case of television, AR could expand the formerly passive media consumption with opportunities for active interaction with the contents – fostering engagement and hopefully, in the long term, increase retention of edutainment content. Research began to explore this concept of AR interaction for television application (cf. [26, 30]). However, prior work has not yet explored the challenges and opportunities for AR in educational television nor investigated the active interaction of the viewer with the content.

In this work, we present the augmented reality *SpaceStation Application*. In collaboration with a German educational television editor and a broadcasting station, we designed video content on the International Space Station (ISS), in which a TV host is presenting and interacting with a model of the ISS. This model can then be interactively explored through the *SpaceStation App* by the viewers as well. Both content and setting are carefully chosen to approximate a realistic scenario of edutainment in television. In a within-subject user study ($N = 31$), we invited participants (21 individuals and five teams of two) to watch two educational videos, one of which is supplemented by the interaction with the *SpaceStation App*. To explore the workload induced through the additional input, we assessed participants' cognitive load with the NASA Task Load Index (NASA-TLX) questionnaire. Additionally, we gathered insights on the users' experience within the app through the User Experience Questionnaire (UEQ). To conclude our investigation, we observed participants' behaviour when watching the videos as well as using the app and requested feedback on their experience.

Our analysis revealed that participants considered the videos as very interesting and informative. They enjoyed exploring the video content in-depth with the *SpaceStation App* and perceived the learning with the app as more creative, innovative, exciting, and enjoyable. However, we noticed that participants were at some points overwhelmed by the amount of information that they perceived through the videos while simultaneously interacting with the app. This was confirmed by the results of the task load questionnaire.

Altogether, we were able to gather in-depth insights into the experience of enriching educational content with augmented reality. We will discuss our findings and outline potential improvements to create adequate AR content for edutainment purposes and to better integrate such an application into the storyline of the television program. Moreover, we will discuss the opportunities and challenges AR brings for different stakeholders in television and state initial ideas for the integration of AR content in television programs.

We conclude our paper with a global discussion how edutainment on other platforms can benefit as well from AR enrichment as presented in this paper.

2 AUGMENTED REALITY

Augmented Reality is defined as virtual objects, overlaying the real environment [19]. AR can be perceived through smartglasses or other head-mounted displays (HMDs; such as the Microsoft HoloLens[2]), or through the camera of handheld devices (HHDs) such as tablets or smartphones. Although HMDs and glasses come with technical advantages over mobile AR, the pervasiveness of smartphones as a mainstream consumer product could be the facilitating factor for people to adopt AR technology [12] in the future.

In the following, we will outline prior research investigating the application of AR in the education sector and

[1]Pokemon Go: https://pokemongolive.com/en/, last accessed January 24th, 2020

[2]Microsoft Hololens: https://www.microsoft.com/en-us/hololens, last accessed January 24th, 2020

summarize what has been tested in the field of (educational) television so far.

AR for Educational Purposes

Almost 20 years ago, Billinghurst [1] already outlined the potential for AR in education, pointing out that its tangible nature can enhance collaboration. In the past decades, AR has been explored in a variety of educational settings. From classroom learning (cf. [4]) over language learning (cf. [7, 11]) to science or STEM learning (cf. [10, 34]), AR has shown to be beneficial on a number of different levels. By extending the real-world environment with digital content, AR technology is particularly interesting for teaching spatial skills.

Prior studies have shown that using AR to learn about complex content, such as components of aircraft turbines, students were able to recall the learned information better compared to learning with traditional materials [16]. In another example from the natural science domain, Sin and Zaman [31] designed an AR textbook to teach students about the solar system. They were able to show that their application was easy to use and supported learners in acquiring knowledge about the solar system.

Furthermore, AR applications have a positive impact on factors that mediate learning, such as students' motivation [6] and task engagement through the immersion with digital content [5, 28]. Research has shown the manifold benefits of AR, making it a promising concept for the enrichment of television programs.

The benefits of AR for learning described in this section go in line with findings from classical learning theories. Through the presentation of contents in multiple modalities, it is often referred to as aligned with the cognitive theory of multimedia learning (CTML) [32].

Augmenting Television

Extending the television program beyond the actual screen and/or to another smart device sounds like an appealing idea. However, prior research has shown that second screen usage, like interacting with a smartphone while watching TV, distracts viewers [21, 22]. Broadcasters and media companies are, thus, experimenting with AR as an extension of TV content, as it keeps the users' focus of attention on the television. Hereby, they are aiming to mitigate the distracting effects of multitasking with a second screen while still extending the television content.

One example is the commercially available smartphone app AugmenTV[3]. It detects the current television program and extends it by displaying additional content and objects around the screen. This technology is currently used for advertisement and entertainment purposes (e.g., storytelling in retail or instant product shopping) but could be extended to educational use as well.

Kim et al. [13] present in their paper a novel hybrid content synchronization scheme for augmented broadcasting services. Aware of second screen apps' distracting nature, they propose an AR approach. Using an AR app, viewers can interact with 3D graphics without leaving the TV screen out of sight. The information is provided on a companion device such as the smartphone. To reduce distraction, AR content is placed in areas in front of the TV screen.

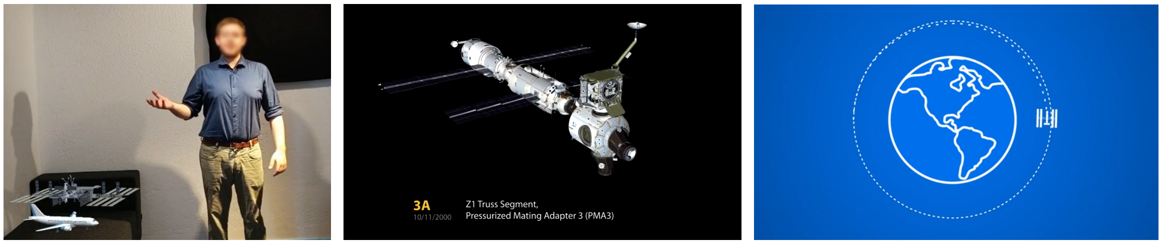
An exemplary use case showing the benefit of AR in television is the work of Ziegler et al. [36], who used AR for the visualization of sign language interpreters in the television program. [36] They developed and tested a HoloLens application for hearing-impaired people with synchronised sign language content to the TV program. They aimed to provide hearing-impaired viewers full accessibility without having to reduce the size of the actual content to fit the interpreter on the same screen. Facing the problem of distraction by shifting between content, the signer overlapped slightly with the TV content so both could be observed directly.

In the domain of television, the implementation of AR into educational programs has been explored as well. For example, Revelle et al. [29] evaluated the use of an AR application during the breaks of a children's TV show called *Electric Agents*. During those breaks children were encouraged to find hidden content to help the characters in the show. The increased interaction with the TV show led to a higher perceived immersion and collaboration [29].

In a recent exploration, Saeghe et al. [30] designed a prototype for viewing AR content through a head-mounted display while watching a BBC nature documentary. For their prototype, they used passive AR content. Viewers were not able to interact with the holograms. For evaluation they focused on experience and engagement. They found that their AR application increased the users' engagement.

Popovici and Vatavu [27] investigated user preferences toward AR applications for television. They reported survey data from 172 participants, showing that the top four television programs they regularly follow are movies, science and technology, documentaries, and educational content. Furthermore, the participants rated 20 potential scenarios according to their perceived value for themselves. Being able to interact with AR content was rated the highest of all scenarios [27]. Since these results are based on questionnaires and hypothetical, research needs to actually evaluate users' responses to such novel concepts.

[3]AugmenTV application: https://augmen.tv/, last accessed January 24th, 2020



(a) (b) (c)

Figure 2: This figure depicts the views used in the final ISS educational videos, showing (a) a moderator view while showing the ISS AR visualization, and (b-c) general graphical visualizations on the ISS as described in the text.

3 SUMMARY AND RESEARCH QUESTIONS

The previous sections emphasized the great amount of research that has been done to investigate the effects of AR use in education and television contexts. It is undoubted that AR can be very beneficial in supporting learning not only in classroom contexts (e.g., [6, 10, 16]), but also in combination with the regular television program. So far, research focused on the technological perspective [13] or the presentation of passive content, such as the sign language interpreter [36]. The latest work of Saeghe et al. [30] presented a prototypical application to extend television with passive AR visualizations. However, they neither investigated participants' view on interactive and engaging AR content, nor did they gather feedback on the additional workload AR could have imposed. Thus, in this work, we aim to address those research gaps by answering the following two main research questions:

RQ1: *How do users perceive/experience the interaction with AR content as a supplement to educational videos?*

RQ2: *Does the interaction with AR content, while viewing educational videos, impose additional workload on the viewer?*

4 SPACESTATION AR APPLICATION

To test our research questions, we designed educational videos and a supplementary AR application to gather insights into the effect of AR on TV-based learning. The scenario depicts a concrete use case of how to combine explanatory videos with an AR app.

We produced two explanatory videos, each about five minutes long. In our videos, we included openly available video materials from the YouTube channels of ESA (e.g., on space debris[4]) and NASA (e.g., the ISS assembly[5] as for example seen in Figure 2b), and extended them by AR visualizations and moderator views. In these views, the host shows and interacts with the AR objects and gives further explanations (for example views from the videos, see Figure 2).

In addition, a smartphone AR app was developed. The app can display the same AR objects as seen in the video and enables the viewer to interact with those. The app is divided into two parts; each part is offering content for its respective video, which is outlined below. Both app parts offer a similar number of interactions, just as the two videos have a similar length and provide roughly the same amount of information. Hereby, we aimed to keep the two videos comparable and reduce any content related effects. The topic of both videos and the app is the International Space Station (ISS). The ISS is a neutral science subject, which can kindle the interest of a diverse group of people. Further, the ISS offers a variety of specifics to formulate questions about, which can then be covered in our explanatory videos. Also, comparable to the aerospace topics of Macciarella et al. [16], many aspects are abstract and can benefit from an AR representation.

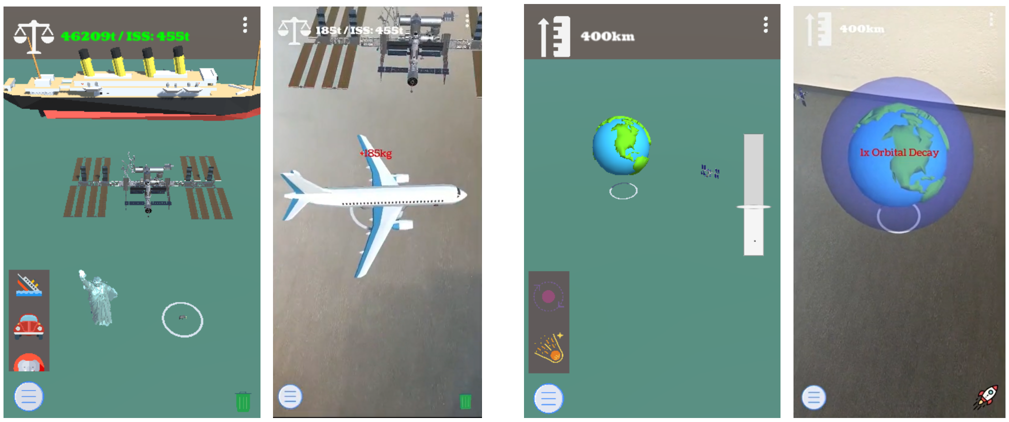
Content

The following section introduces the two topical parts of the videos and the representation in the AR application. To confirm the content and structure of the educational videos, we consulted an editor of a German broadcasting company. This editor is currently working in educational television and provided us with essential feedback. We revised the videos according to his feedback and let him confirm the quality of the final videos. The videos were produced in German for a German-speaking audience.

Part I - ISS in Relation to Earth Objects. The first video provides basic information about the ISS and includes (1) the historical background and structure of the ISS, (2) how astronauts survive there and finally, in a vivid comparison, (3) the size and weight of the ISS are described. In the first part of the app, the user sees the ISS in relation to objects from earth (e.g., vehicles, animals, landmarks) to get an impression of its size (cf. Figure 3b). The app presents a list of objects to

[4]YouTube video on soace debris: https://www.youtube.com/watch?v=wkJ3vEUiC9g, last accessed January 24th, 2020

[5]YouTube video on the ISS assembly: https://www.youtube.com/watch?v=0WKOuaXd0oE, last accessed January 24th, 2020



(a) (b) (c) (d)

Figure 3: This figure shows the designs for the content part I ("ISS in Relation to Earth Objects") as (a) the initial design for the app interface as well as (b) the final implementation in the AR *SpaceStation App*. Furthermore, it shows the realization of content part II ("ISS and its Orbit"), as (c) the initial design and (d) final implementation.

choose from, and the user can arrange those objects virtually next to the ISS. The objects and the ISS can be scaled together keeping the real size ratio. Moreover, the *SpaceStation App* shows the weight of the objects as well as the weight of the ISS. In this first part, we make use of AR's potential to re-size, re-position, and compare different objects in a virtual space.

Part II - ISS and its Orbit. Extending the content of part I, part II provides more explanation on the orbit of the ISS around Earth. The video answers the questions of (1) how fast the ISS moves, (2) what the loss of altitude (called "Orbital Decay") means, and (3) why the ISS needs to keep the height of 400km. To explain this, a 3D model of the earth is visible including its atmosphere. The *SpaceStation App* shows the ISS orbiting the earth, displaying the distance/height that it has from earth. The video further shows the concept of Orbital Decay, explaining the process of two orbital objects (i.e., the earth and the ISS) approaching each other. In the respective video, this process is visualized by the ISS losing height, while the user can accelerate or decelerate the process through the menu of the app. Visual warnings are presented when the ISS reaches critically low altitudes, requiring the user to push a 'reboost'-button so that the ISS can increase its altitude again.

Technical Background and Implementation

For the development of the app, we used the game engine Unity[6]. To implement the AR content, we used the package *ARFoundation*. *ARFoundation* offers the advantage that the code can be deployed on both Android and iOS devices. In addition, the implementation of some basic features of AR in *ARFoundation* is greatly streamlined and simplified compared to *ARCore* or *ARKit*.

AR Interaction. The initial placement of all objects can be controlled by tilting the device. The object can only be placed on a detected surface. During placement, the object will always rotate toward the camera and stay at the center of where the camera is facing, if there is a detected surface at the center. Once the initial placement is confirmed, the object can be moved again by dragging it with one finger. The selected object can also be rotated with two fingers. If no object is currently selected, all placed objects can be scaled simultaneously by using the pinch gesture, while keeping their real-world proportions.

There are two menus accessible in the app. One menu can be opened in the upper right corner of the screen (three little buttons as an icon, see Figure 3). It contains two buttons: (1) reload scene (removes objects and resets tracking) and (2) hide user interface (feature also used to record the videos of the host interacting with the AR objects). The second menu offers options specific to the selected object. In part one, this special menu consists of the list of items that can be added to the scene (see Figure 3a). In part two, the Orbital Decay can be adjusted, and the slider can be toggled to change the height of the orbiting ISS (see Figure 3c).

Video Production. The videos consist of an alternation of moderator scenes as well as animations and videos with continued commentary of the moderator (cf. Figure 2). In the moderator scenes, the host in the studio explains a topic, while a 3D model from the app, matching the subject, is placed next to him. The moderator scenes were recorded

[6]Unity Application: https://unity.com/, last accessed Januar 24th, 2020

with a Pixel 3 smartphone. In order to include the 3D models into the video during recording, the app was opened and directed at the presenter, the models where placed, and the screen was captured. Audio was recorded with a second device. Unfortunately, some image quality was lost by this method so that the resolution of the video had to be reduced to 720p to provide an acceptable viewing experience. Both videos last about five minutes. We created new video content for this project to be able to have 3D objects from the *SpaceStation App* that are also visible in the video, to create a further connection between app and video.

5 EVALUATION

With our study, we aim to gain a deeper understanding of the effects of having a supplemental AR application for educational videos on the user experience of our participants. For the user study setup described in the following, we routinely requested ethics approval from the ethics committee of our university, who accepted our proposal.

Experimental Design

The study follows a within-subject design and contains quantitative as well as qualitative evaluation parts. As our independent variables, the participants watched the video either while using the *SpaceStation App* or without using it. The videos themselves were the same for both conditions and the order of the videos was controlled according to Figure 4 to mitigate any effects of the video contents on the user experience and workload while watching/interacting with the content. As confounding variables, we also noted any usability issues and problems of understanding, regarding content and usage of the application.

Our study sample consisted of 31 participants, of which 21 took part in the user study on their own and 10 took part in teams of two. In the paired setting, the participants knew each other (friends or family members) and interacted with the *SpaceStation App* on individual devices. With this variation of our study setup, we aimed to explore the feasibility of our application in a multi-user setting.

Procedure

At the beginning of our user study, the participants filled out a short questionnaire about demographic information (gender, age, profession). We informed them about our anonymization and data protection policy in line with the GDPR. After giving informed consent, the participants specified their prior knowledge on the ISS on a 7-point Likert (1 = no knowledge, 7 = expert knowledge) and their previous experience with AR applications (1 = never used AR apps, 7 = frequently used AR apps). Subsequently, we introduced participants to an AR application called *Slither*, based on a tutorial for

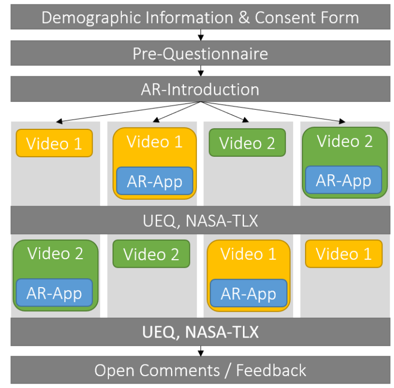


Figure 4: The study procedure as applied in this user study.

ARCore[7]. In Slither, participants played the game "Snake" in an AR setting. To move the snake and find the food pellets, the user has to move the smartphone accordingly. Thus, the app introduced an important aspect of smartphone AR: The scanning of surfaces by panning the camera over a chosen surface. By including this game, we intended to mitigate the influence of the novelty effect when using currently unfamiliar technology in a user study [35].

Before the actual study started, we informed the participants about their task of learning about the ISS. They were asked to engage with the contents and try to understand the topic. After the introduction, the participants watched the two explanatory videos appearing in alternating order, as shown in Figure 4. While watching one of the two videos, each user interacted with the *SpaceStation App*. The contents did not build up on each other, which allowed us to randomize the order of the videos. For the AR conditions, we included a prompt at the end of each video, inviting the participants to interact with the *SpaceStation App*. However, to observe the users' natural behaviour, we did not restrict the participants in using the application while the video was playing.

Finally, after each video, the participant filled in two questionnaires, the *NASA Task Load Index* (NASA-TLX) [9], and the *User Experience Questionnaire* (UEQ). At the end of the two questionnaires, a text input field enabled the participants to leave individual notes and comments on their experience. The questionnaires were filled in electronically, on the same device that was used for the *SpaceStation App*. In total, the participation in our study required between 20 and 30 minutes, depending on the depth of interaction with the application. We thanked the participants for their time and as compensation for the participation they received sweets.

[7]ARCore Tutorial: https://codelabs.developers.google.com/codelabs/arcore-intro/, last accessed January 24th, 2020

Apparatus & Setup

The study took place in a lab environment designed as a living room to make participants feel comfortable and mimic a common usage situation. The users interacted with a Samsung Galaxy Tab S4[8], a tablet with a 10.5-inch display. In the paired setting, the second participant received a Google Pixel 3[9] smartphone with a 5.5-inch display. We decided to use hand-held devices for the AR presentation instead of glasses in order to portrait a realistic use case situation in which family members or friends view the television program together with already available hardware, also allowing multi-user settings in the users' living room.

The videos played on a 50-inch TV, and the users sat on a sofa while watching the video and interacting with the *SpaceStation App*. The instructor waited close by, to answer any emerging questions without influencing the user or interfering with the interaction.

Measures

We presented the participants with two questionnaires, UEQ and NASA-TLX, and observed the participants in their interaction with the *SpaceStation App*.

The **NASA Task Load Index (NASA-TLX)** is an instrument to assess the perceived workload that a user is experiencing while performing a task or interacting with an application [9]. It consists of six facets presented in the form of a 7-point Likert item (1 = low workload, 7= high workload), including (1) mental demand, (2) physical demand, (3) temporal demand, (4) performance, (5) effort, and (6) frustration. We applied the NASA-TLX to gain insight into potential usability problems of our application and to get an impression of the workload that is induced by the additional interaction with the *SpaceStation App*.

The **User Experience Questionnaire (UEQ)** is a questionnaire containing 26 pairs of opposing items. We applied it to assess the overall user experience during the videos and interaction with the *SpaceStation App*. For each word pair of the UEQ, the participants had to indicate which word is more fitting to describe their user experience on a 7-point Likert scale (e.g., 1 = boring, 7 = exciting). The 26 items contained word-pairs from six overall topics: (1) attractiveness, (2) perspicuity, (3) efficiency, (4) dependability, (5) stimulation, and (6) novelty. While (1) is a dimension of valence, (2-4) assess the pragmatic quality and (5-6) the hedonic quality. The UEQ is available in several languages and the full set of items can be downloaded from the official UEQ webpage[10]. Furthermore, the UEQ has been evaluated in terms of construct validity [14, 15].

Both questionnaires were presented in German after both study parts, to assess the user experience and workload for the video and/or the interaction with the *SpaceStation App* , respectively.

Additional to the two questionnaires described above, we included an open comment section in our questionnaires, aiming to collect further comments and opinions regarding our application and user study. During the study, the instructor observed the participants' interaction to uncover difficulties in the usage of the application, taking notes on any unexpected action, verbal comments, or feedback. Furthermore, at the end of each study, the instructor directly asked each participant verbally for any further feedback.

Sample

Thirty-one participants took part in our user study (12 female, 19 male). We recruited a diverse set of participants, age 15 to 65 ($M = 31.32$, $SD = 12.75$) through personal contacts and internal mailing lists. For participants who were not of legal age (eighteen in Germany) we additionally requested written consent from their parents or legal guardians. As a requirement for our study, we excluded participants with in-depth knowledge of the ISS, to target a low prior knowledge level, as well as participants without experience in using a smartphone or tablet. Participants reported a low-medium prior knowledge on the ISS ($M = 3.06$, $SD = 1.32$) and low-no experience with AR applications ($M = 3.1$, $SD = 1.74$). Twelve participants were university students, four high-school students, and fifteen had full-time jobs in the broadcasting and media sector.

6 RESULTS

In the paired setting with two participants, both interacted with their individual device, since the app does not include any collaborative features, and rated their experience and workload independently. However, in these situations, we can not neglect the potential impact of the study partner on the workload and user experience. To decide whether or not to analyze the data as one large or two separate groups, we performed a Mann-Whitney-U Test. This test was used to uncover potential differences in the UEQ and NASA-TLX scores between the two independent groups, which would require us to process the data separately. The comparisons yielded no significant differences in any facet. Due to our small and unevenly distributed sample size between the groups, we decided not to analyze the quantitative data separately. We do,

[8]Samsung Tablet used in study: https://www.samsung.com/tablets/galaxy-tab-s4/,last accessed January 24th, 2020

[9]Google Phone used in study: https://store.google.com/product/pixel_3, last accessed January 24th, 2020

[10]User Experience Questionnaire: https://www.ueq-online.org/, last accessed January 24th, 2020

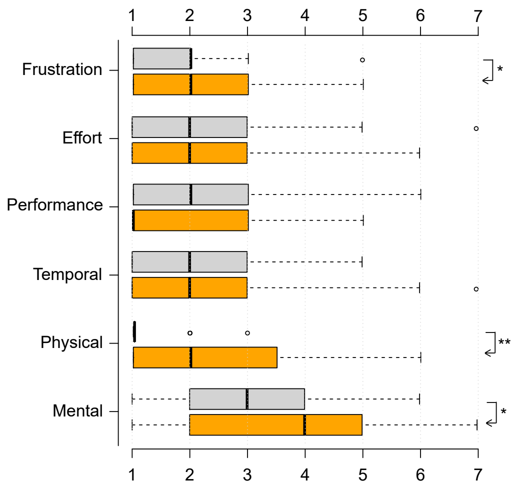


Figure 5: The results of the NASA-TLX. The ratings for watching the videos with the *SpaceStation App* are depicted in orange, without the app in light grey. The significant differences found using the Wilcoxon Signed Rank Test are marked with $p < .05$ (*) and $p < .001$ (**).

however, investigate the subjective feedback and interview reports in regard to potential problems of the application for a multi-user scenario, to make sure that the setting did not negatively affect the participants' workload or experience.

NASA-TLX

For the NASA-TLX evaluation ($N = 31$), we started by analyzing the data for normal distribution. A Shapiro-Wilk test revealed a deviation from normality ($p < .05$) for the NASA-TLX values after watching the videos with and without the *SpaceStation App* . Thus, we continued our analysis with a non-parametric Wilcoxon test, showing significant differences in the facets mental workload ($p < .05$, $Z = -2.226$), physical workload ($p < .001$, $Z = -3.909$), and frustration ($p < .05$, $Z = -2.352$). The descriptive data (see Figure 5) shows for all three significant deviations, with the no_app condition having lower NASA-TLX scores (mental workload: $M_{\text{App}} = 3.83$, $M_{\text{NoApp}} = 2.77$; physical workload: $M_{\text{App}} = 2.71$, $M_{\text{NoApp}} = 1.1$; frustration: $M_{\text{App}} = 2.29$, $M_{\text{NoApp}} = 1.65$). Thus, the data indicates that enabling the users to interact with the app in addition to watching the video does induce a higher mental and physical workload as well as a higher frustration level compared to watching the video without the app.

UEQ

Similar to the NASA-TLX data, the UEQ results showed a deviation from normality in a Shapiro-Wilk test ($p < .05$). We followed up with the non-parametrical Wilcoxon test, revealing a significant difference in 14 of the 26 UEQ facets. Besides others, viewing the video and using the supplementary *SpaceStation App* was considered as more creative, exciting, supportive, innovative, and easier to learn. However, the results also showed that the app was considered less well structured, slightly more confusing, and less compliant with the users' expectations of the system (for full comparison see Figure 6). We will further discuss these findings in the subsequent section.

Open Comments & Observations

In the open comment field of the questionnaire, participants stated to find the videos well made (P3, P12), informative (P12), and the topic interesting (P3, P15)[11]. The highschool students further noted that they would appreciate such apps to become part of their daily classes.

They perceived the *SpaceStation App* as a "nice, interactive addition [to the video]" (P5, similarly repeated by P9), enjoyed using it (P20), and highlighted individual features of the app, such as the tooltips (P19). However, participants also noted having had difficulties in multitasking between the video and the *SpaceStation App* , saying: "At the beginning, it was difficult to follow the video, because I was busy interacting with the app" (P22 similarly repeated by P6). One participant stated that he would like to be able to adapt the speed of the video and to closer couple the presentation in the video with the application (P9). Our observations supported this finding. The speed of the video often did not allow interaction with the app before the end of the video. Since the speed of the television program is fixed, broadcasters and editors need to consider additional time for interaction when creating TV content with AR supplements. We received no feedback indicating a negative effect of the multi-user setting on the users' experience and workload.

7 LIMITATIONS

The aim of our user study with the *SpaceStation App* was to discover how viewers perceive AR-supported edutainment content. Besides enjoyment and acceptance, we also wanted to assess if the users perceive any distraction or additional workload induced by the app.

We did not measure the actual learning outcome, since it would be strongly influenced by our study conditions, prior knowledge, and the extent of the novelty effect. However, we focused on assessing moderating factors such as enjoyment and motivation through the UEQ. Further studies in a more controlled setting are needed to measure the differences in the learning process and retention.

[11]Note: For the reporting of the qualitative results, all comments and quotes were translated to English from the participant's native language.

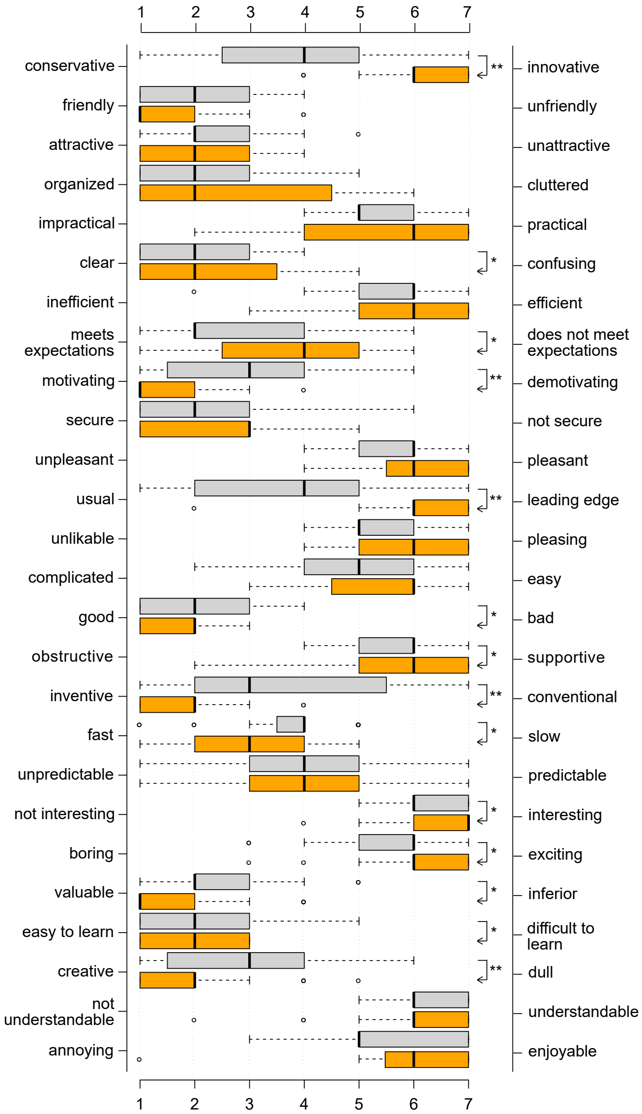


Figure 6: The results of the UEQ. The ratings for watching the videos with the *SpaceStation App* are depicted in orange, while the video without supplementary app is shown in light grey. The significant differences found using the Wilcoxon Signed Rank Test are marked with $p < .05$ (*) and $p < .001$ (**).

In our study, we started the app and the videos manually, which was appropriate for our setting. It would have been nice to synchronize the app and the edutainmant content, for example using HbbTV. In a real use-case, such synchronization would be necessary to lower the viewers' effort. Furthermore, we could not capture the moderator's interactions with the AR objects and replay them in the app. It might have a positive effect on the cognitive workload if the participants in the study would not be able to control the AR objects at the same time as the moderator.

Furthermore, the participants had to set up the AR environment themselves (i.e., they had to do the surface detection and object placing on their own). If users changed the setup while they watched the video, this may have caused distraction. In a controlled setting, the surface detection could have been done through the use of markers. For the automated surface detection to work in every living room, Unity's *Project MARS*[12] should be considered for future work. This feature allows for automatic detection of the TV's position as well as a coffee table to place the virtual scene onto.

8 DISCUSSION

In our study, we were able to gather valuable insights into the users' perception of AR contents and its integration into the television program. We discuss those results according to our two research questions and our further observations.

User Experience (RQ1). In general, participants enjoyed interacting with the *SpaceStation App*. As reported in the Results section, they described it as a useful addition to the bare consumption of the videos. Thus, novelty could have still affected the UEQ scores, especially the higher scores for the *SpaceStation App* in the hedonic quality of the UEQ (excitement, interest, innovation, etc.). Thus, only a long term evaluation on the use of an AR app for a television program could reveal if the shown effects decrease by the time the novelty of the technology decreases. However, due to the diversity of potential interactions that an AR application would bring to educational television, we are confident that it could spark motivation and engagement over a longer period of time.

Workload and Potential for Distraction (RQ2). The comparison of the NASA-TLX scores revealed a slight increase in workload during the use of the *SpaceStation App* . The increase of the physical workload can be explained by the physical interaction with the AR objects and still remains at a very low level of effort. The mental workload, however, can be a result of the multitasking between video and app interaction, as described by P6 and P22 and confirmed by observing many other participants. In our study, the participants had the opportunity to interact with the app while watching the video. To prevent the necessity of multitasking and, thus, decrease the workload, we recommend designing learning contents with designated moments for interaction. In our current design, users can interact with the content during or at the end of the videos. If they chose to interact already during the video, the users reported problems with

[12]Unity project MARS: https://unity.com/unity/features/mars, last accessed January 24th, 2020

multitasking and additional workload, because the video did not support interaction breaks. We envision future designs to have shorter but more frequent possibilities for interaction, which are guided by the TV host and coupled to breaks in the explanatory videos. On the one hand, this would leave the viewers with less flexibility and control over the interaction, but on the other hand enables them to focus on the videos and the interaction individually. These proposed changes go in line with the participants' wish to adapt the speed of the video. This would not be necessary in a setting with designated interaction breaks.

Sustainability - A Stakeholders' Perspective. Another aspect that has to be discussed is the suitability of this technology for a wider distribution. On the one hand, content producers, moderators, and viewers have to accept the technology with its potential and challenges. On the other hand, an appropriate way of distribution has to be found for the application to become more widely used.

Regarding the acceptance of the technology, we can say so far that viewers liked the approach and would like to see more content enriched by AR apps. Especially the younger participants (highschool and university students) were impressed by the technology and would appreciate an integration into their courses as well. Therefore, we consider the production of AR supported edutainment content a promising way to target a younger audience. The integration of AR into the television program would come with a set of changes and obstacles for the content producers and hosts. A change in the actual production chain would be necessary to produce useful AR-supported content. Today, graphical augmentation is part of the post-production. To enable the moderator to interact with the AR content and to provide the viewers with AR objects, the graphic workflow has to be transferred from post- to pre-production. AR objects should be treated like requisites. Using game engines like *Unity* enables graphic designers to program the behaviour of virtual content to enhance interactivity and real-time feedback. However, these non-linear workflows are very new to the editorial staff and the whole production team. Interacting with AR objects might also be challenging for moderators. Future work has to evaluate if hosts would accept such technologies and how we can design applications to support the creation and live interaction of contents from the host perspective. However, this new technology comes with many opportunities. Apart from the benefits discussed in this paper, AR apps may be a solution for broadcasters to break out of the linear content distribution. Viewers might control through the AR app what they want to see on TV. For example, based on the component of the *ISS* that the user is looking at or interacting with in AR, the fitting content can be presented.

Interactive AR App's - Enrichment for Edutainment. An AR app, similar to the one presented in this paper, might not only be beneficial for broadcaster but also for other content providers like TED or YouTube. However, they are likely to face other challenges and encounter other opportunities than they occur in the case of broadcasting. For instance, changing the platform comes with changes in viewer habits. More than 70% of YouTube's content is viewed on mobile devices [20]. Following the approach of the *SpaceStation App* and presenting video and AR content in parallel would be challenging due to limited screen space. Thus, YouTube's first AR extension implements a split-screen, showing the augmented content below the explanatory video [24]. In case of mobile usage, it might be more interesting to edit the edutainment content completely in AR. However, streaming platforms like Netflix or the German ARD Mediathek, which provide a flexible web infrastructure, may benefit strongly from an AR app.

9 CONCLUSION & FUTURE WORK

In this work, we designed the *SpaceStation* augmented reality application, built to provide a new way of interacting with otherwise passive television content. This research is our first step toward integrating AR to enrich edutainment programs. Through our user study ($N = 31$), we were able to highlight the importance of implementing designated breaks for viewers to interact with supplementary AR content, thus, mitigating the additional workload that it imposes on the viewer. Although we could not find any negative effects of the multi-user setting on user experience or workload, future work is needed to take an in-depth view on this use case.

On the positive side, our findings show the users' openness toward this technology and their predominantly positive experiences with the *SpaceStation App* . Especially for the application in STEM edutainment television, we see great potential for augmented reality, due to its great variety of visualization and interaction possibilities. In contrast to using AR for passive information presentation (cf. [30]), the *SpaceStation* concept enables the moderator to flexibly interact with the AR model.

For future research, we envision that the AR interaction could even become a tool to implicitly collect user feedback, i.e., uncovering the most interesting topic or detecting problems of understanding through interacting with the content. Through the aggregation of the users' interaction patterns, we could provide the host with the opportunity to react to the users in-time, creating a bi-directional feedback loop. Since AR has already proven its benefits for education and other contexts, we hope to encourage researchers and practitioners to create and evaluate AR applications for edutainment context in other media formats as well.

REFERENCES

[1] Mark Billinghurst. 2002. Augmented reality in education. *New horizons for learning* 12, 5 (2002), 1–5. https://doi.org/10.1016/j.sbspro.2012.06.654

[2] Gary Bitter and Allen Corral. 2014. The Pedagogical Potential of Augmented Reality Apps. *International Journal of Engineering Science Invention* 3, 10 (2014), 13–17. https://doi.org/10.6084/m9.figshare.1295969.v1

[3] TED Blog. 2012. TED reaches its billionth video view! https://blog.ted.com/ted-reaches-its-billionth-video-view/ , Last accessed January 17th, 2020.

[4] Sébastien Cuendet, Quentin Bonnard, Son Do-Lenh, and Pierre Dillenbourg. 2013. Designing augmented reality for the classroom. *Computers & Education* 68 (2013), 557–569. https://doi.org/10.1016/j.compedu.2013.02.015

[5] Chris Dede. 2009. Immersive interfaces for engagement and learning. *science* 323, 5910 (2009), 66–69. https://doi.org/10.1126/science.1167311

[6] Ángela Di Serio, María Blanca Ibáñez, and Carlos Delgado Kloos. 2013. Impact of an augmented reality system on students' motivation for a visual art course. *Computers & Education* 68 (2013), 586–596. https://doi.org/10.1016/j.compedu.2012.03.002

[7] Min Fan, Uddipana Baishya, Elgin-Skye Mclaren, Alissa N. Antle, Shubhra Sarker, and Amal Vincent. 2018. Block Talks: A Tangible and Augmented Reality Toolkit for Children to Learn Sentence Construction. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, Article Paper LBW056, 6 pages. https://doi.org/10.1145/3170427.3188576

[8] Funk. 2016. Hey, wir sind funk. https://www.funk.net/funk , Last accessed January 17th, 2020.

[9] Sandra G Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (2006), 904–908. https://doi.org/10.1177/154193120605000909

[10] María Blanca Ibáñez, Ángela Di Serio, Diego Villarán, and Carlos Delgado Kloos. 2014. Experimenting with electromagnetism using augmented reality: Impact on flow student experience and educational effectiveness. *Computers & Education* 71 (2014), 1–13. https://doi.org/10.1016/j.compedu.2013.09.004

[11] Adam Ibrahim, Brandon Huynh, Jonathan Downey, Tobias Höllerer, Dorothy Chun, and John O'donovan. 2018. Arbis pictus: A study of vocabulary learning with augmented reality. *IEEE transactions on visualization and computer graphics* 24, 11 (2018), 2867–2874. https://doi.org/10.1109/TVCG.2018.2868568

[12] IQUII. 2018. From Augmented Reality to Mobile Reality: it's the smartphone that drives augmented reality. https://medium.com/iquii/from-augmented-reality-to-mobile-reality-its-the-smartphone-that-drives-augmented-reality-31ae65b44958 , Last accessed January 17th, 2020.

[13] Soonchoul Kim, Bumsuk Choi, Youngho Jeong, Jinwoo Hong, and Kyuheon Kim. 2014. Novel hybrid content synchronization scheme for augmented broadcasting services. *ETRI Journal* 36, 5 (2014), 791–797. https://doi.org/10.4218/etrij.14.0114.0346

[14] Bettina Laugwitz, Theo Held, and Martin Schrepp. 2008. Construction and Evaluation of a User Experience Questionnaire. In *HCI and Usability for Education and Work*, Andreas Holzinger (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 63–76. https://doi.org/10.1007/978-3-540-89350-9_6

[15] Bettina Laugwitz, Ulf Schubert, Waltraud Ilmberger, Nina Tamm, Theo Held, and Martin Schrepp. 2009. Subjektive Benutzerzufriedenheit quantitativ erfassen: Erfahrungen mit dem User Experience Questionnaire UEQ. *Tagungsband UP09* (2009).

[16] Nickolas D Macchiarella and Dennis A Vincenzi. 2004. Augmented reality in a learning paradigm for flight aerospace maintenance training. In *The 23rd Digital Avionics Systems Conference (IEEE)*, Vol. 1. IEEE, 5–D. https://doi.org/10.1109/DASC.2004.1391342

[17] mediakix. 2019. The 16 Most Popular Types Of YouTube Videos. https://mediakix.com/blog/most-popular-youtube-videos/ , Last accessed January 17th, 2020.

[18] Lothar Mikos. 2016. Digital media platforms and the use of TV content: Binge watching and video-on-demand in Germany. *Media and Communication* 4, 3 (2016), 154–161. https://doi.org/10.17645/mac.v4i3.542

[19] Paul Milgram and Fumio Kishino. 1994. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems* 77, 12 (1994), 1321–1329.

[20] Maryam Mohsin. 2019. 10 Youtube Statistics 2020 Every Marketer Should Know. https://www.oberlo.com/blog/youtube-statistics

[21] Timothy Neate, Michael Evans, and Matt Jones. 2016. Designing visual complexity for dual-screen media. *Conference on Human Factors in Computing Systems - Proceedings* May (2016), 475–486. https://doi.org/10.1145/2858036.2858112

[22] Timothy Neate, Matt Jones, and Michael Evans. 2015. Mediating attention for second screen companion content. *Conference on Human Factors in Computing Systems - Proceedings* 2015-April, July (2015), 3103–3106. https://doi.org/10.1145/2702123.2702278

[23] Stefan Niggemeier. 2016. "funk": ARD und ZDF machen Youtuber fett | Übermedien. https://uebermedien.de/8308/funk-ard-und-zdf-machen-youtuber-fett/ , Last accessed January 17th, 2020.

[24] Tommy Palladino. 2019. YouTube Unleashes First AR Makeup Try-On Tool with Mac Cosmetics. https://mobile-ar.reality.news/news/youtube-unleashes-first-ar-makeup-try-tool-with-mac-cosmetics-0203389/

[25] Nikolaos Pellas, Panagiotis Fotaris, Ioannis Kazanidis, and David Wells. 2019. Augmenting the learning experience in primary and secondary school education: a systematic review of recent trends in augmented reality game-based learning. *Virtual Reality* 23, 4 (2019), 329–346. https://doi.org/10.1007/s10055-018-0347-2

[26] Irina Popovici and Radu-Daniel Vatavu. 2019. Towards Visual Augmentation of the Television Watching Experience: Manifesto and Agenda. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video*. ACM, 199–204. https://doi.org/10.1145/3317697.3325121

[27] Irina Popovici and Radu-Daniel Vatavu. 2019. Understanding Users' Preferences for Augmented Reality Television. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 269–278. https://doi.org/10.1109/ISMAR.2019.00024

[28] Iulian Radu. 2014. Augmented reality in education: A meta-review and cross-media analysis. *Personal and Ubiquitous Computing* 18, 6 (2014), 1533–1543. https://doi.org/10.1007/s00779-013-0747-y

[29] Glenda Revelle, Emily Reardon, Kristin Cook, Lori Takeuchi, Rafael Ballagas, Koichi Mori, Hiroshi Horii, Hayes Raffle, Maria Sandberg, and Mirjana Spasojevic. 2015. Electric Agents: Combining Collaborative Mobile Augmented Reality andWeb-Based Video to Reinvent Interactive Television. *Computers in Entertainment (CIE) - Theoretical and Practical Computer Applications in Entertainment* 12, 3 (2015), 1–21. https://doi.org/10.1145/2702109.2633413

[30] Pejman Saeghe, Sarah Clinch, Bruce Weir, Maxine Glancy, Vinoba Vinayagamoorthy, Ollie Pattinson, Stephen Robert Pettifer, and Robert Stevens. 2019. Augmenting Television With Augmented Reality. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video*. ACM, 255–261. https://doi.org/10.1145/3317697.3325129

[31] Aw Kien Sin and Halimah Badioze Zaman. 2010. Live Solar System (LSS): Evaluation of an Augmented Reality book-based educational tool. In *2010 International Symposium on Information Technology*, Vol. 1. IEEE, 1–6. https://doi.org/10.1109/ITSIM.2010.5561320

[32] Peter Sommerauer and Oliver Müller. 2018. Augmented reality for teaching and learning–a literature review on theoretical and empirical foundations. *In Proceedings of the 26th European Conference on Information Systems (ECIS2018), Portsmouth, UK, 2018* (2018).

[33] Erin M Steffes and Philippe Duverger. 2012. Edutainment with videos and its positive effect on long term memory. *Journal for Advancement of Marketing Education* 20, 1 (2012), 1–10.

[34] Martin P Strzys, Sebastian Kapp, Michael Thees, Pascal Klein, Paul Lukowicz, Pascal Knierim, Albrecht Schmidt, and Jochen Kuhn. 2018. Physics holo.lab learning experience: Using smartglasses for augmented reality labwork to foster the concepts of heat conduction. *European Journal of Physics* 39, 3 (2018), 035703. https://doi.org/10.1088/1361-6404/aaa8fb

[35] John D Wells, Damon E Campbell, Joseph S Valacich, and Mauricio Featherman. 2010. The Effect of Perceived Novelty on the Adoption of Information Technology Innovations: A Risk/Reward Perspective. *Decision Sciences* 41, 4 (2010), 813–843. https://doi.org/10.1111/j.1540-5915.2010.00292.x

[36] Christoph Ziegler, Richard Schäffer, Vinoba Vinayagamoorthy, Maxine Glancy, Paul Debenham, and Alastair Bruce. 2018. Personalising the TV experience with augmented reality technology: Synchronised sign language interpretation. *TVX 2018 - Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video* (2018), 179–184. https://doi.org/10.1145/3210825.3213562