

Lead me gently: Facilitating knowledge gain through attention-aware ambient learning displays

Citation for published version (APA):

Börner, D., Kalz, M., & Specht, M. (2014). Lead me gently: Facilitating knowledge gain through attention-aware ambient learning displays. *Computers and Education*, 78, 10-19. <https://doi.org/10.1016/j.compedu.2014.04.017>

DOI:

[10.1016/j.compedu.2014.04.017](https://doi.org/10.1016/j.compedu.2014.04.017)

Document status and date:

Published: 01/09/2014

Document Version:

Peer reviewed version

Document license:

CC BY-NC-ND

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

<https://www.ou.nl/taverne-agreement>

Take down policy

If you believe that this document breaches copyright please contact us at:

pure-support@ou.nl

providing details and we will investigate your claim.

Downloaded from <https://research.ou.nl/> on date: 30 Dez. 2021

Open Universiteit
www.ou.nl



Abstract: This empirical study reports an intervention to investigate identified research challenges on the evaluation and use of ambient displays in a learning context with the objective to gain insights into the interplay between display design, user attention, and knowledge acquisition. The main research questions were whether an attention-aware display design can capture the user's focus of attention and whether this has an influence on the knowledge gain. A display prototype corresponding to the main ambient display characteristics was designed, applied in a controlled authentic setting, and evaluated accordingly. The prototype presented information and guidelines for first responders in emergency situations, especially in cases of cardiac arrest. The prototype was enhanced with a custom-built sensor to measure user attention and trigger interruptive notifications. The study was conducted among 52 employees working at a university campus. Using an experimental research design, a treatment group exposed to an attention-aware display design was compared to a control group. The results provide evidence that such a display design can attract and retain attention in such a way that the acquisition of knowledge (i.e. the comprehension of the presented information) is effectively facilitated.

Keywords: ambient learning displays; empirical study; ubiquitous learning support; knowledge acquisition; user attention

1. Introduction

The concept of ambient learning displays is the author's approach to design and implement ambient displays in a learning context (Börner, Kalz, & Specht, 2013a). The idea is to provide situated support of learning scenarios in ubiquitous learning environments by enabling learners to view, access, and interact with contextualised digital content presented in an ambient way (Börner, Kalz, & Specht, 2011). This empirical study is the second in the research and development of ambient learning displays. A first empirical study evaluated the effect of different ambient display designs to initiate learning in an environmental education context (Börner, Kalz, & Specht, 2013b). Based on the design dimensions identified by Pousman & Stasko (2006) in their taxonomy of ambient information systems, the designs were varied on their level of notification and their representational fidelity. The notification level depicts the degree of user interruption ranging from change blind to interruptive. The representational fidelity describes how the data is encoded from indexical to symbolic representations. The hypothesis that using a design with interruptive notification and symbolic representation results in significantly larger learning outcome (comprised of awareness, confidence, knowledge, concern, and attitude) could not be supported. The group exposed to this design still had the largest learning outcome after all and indicated that the

design increased their awareness, but the group exposed to the change blind and indexical design showed the largest knowledge gain. One of the main reasons for the lack of evidence for the hypothesis as well as the major limitation of the study was the experimental setting with high authenticity spanning between unobtrusiveness and user attention.

This challenge is in line with earlier communicated issues regarding the evaluation of mobile and pervasive technologies. Kaikkonen, Kekäläinen, Cankar, Kallio, & Kankainen (2005) discussed the challenges between laboratory studies and field studies for usability evaluations of mobile technology. While lab settings offer a context without the danger of uncontrollable external variables, they have also been criticised as having a very low ecological validity. On the other hand, field settings suffer from many external variables that can influence the results of an experiment while being highly authentic and therefore offering the best ecological validity possible. Alt, Schneegaß, Schmidt, Müller, & Memarovic (2012) discussed the same problem for the evaluation of public displays. The authors summarise that most studies have either focused on internal validity through lab studies or ecological validity with field studies. The core question is thus how to evaluate in realistic settings controlling confounding variables. In an evaluation paradigm spectrum ranging from controlled lab studies to a realistic field studies, the next step would be to strive for something in between – a controlled but authentic setting emulating the dynamics of the real world. Another research challenge derived was how to retain the user interest over time. In this context the study yielded the need for direct feedback and the use of motivational designs, such as gamification (Werbach & Hunter, 2012).

Additionally the influence of the different design variations on the user attention towards the display was examined in another yet unpublished study. Focusing solely on interaction aspects, quantitative measures gathered with non-intrusive evaluation techniques were combined with qualitative measurement of user perception and comprehension to provide a more holistic view of user attention. Again the results did not show any significant effects of different display designs, but suggest designing displays with graspable information and gentle notification to effectively capture and retain user attention. As major limitation the used measurement did not reliably capture single user attention in this

uncontrolled setting. Furthermore from a qualitative perspective the attention-aware design of the displays was challenged by the tensions between noticeability, disruptiveness, and comprehension. Hence, in this empirical study another ambient display prototype was designed accordingly, applied in a controlled experimental setting, and evaluated in a learning context.

1.1 Background

The characterisations of ambient displays are diverse and multifaceted, still mostly building upon the definition by Wisneski et al. (1998). In general, ambient displays are characterised as informative appliances that are embedded into the physical environment. The embedding is supported and fostered by an unobtrusive and peripheral design using subtle communication methods mainly out of the focus of attention. Thereby the displays are able to move from the periphery to the focus of attention and back. These characteristics are complemented by several general requirements, such as to be glanceable and pre-attentively comprehensible as well as not distracting nor demanding attention. The characteristics and requirements of ambient displays also determine their application context. In their work on peripheral displays Matthews, Dey, Mankoff, Carter, & Rattenbury (2004) described the main ability to allow “a person to be aware of information while she is attending to some other primary task or activity.” Focusing on the user attention towards a display the authors presented a toolkit for managing attention through display design. Based on existing models they distinguished three types of attention: inattention, divided attention, and focused attention. Users might ignore peripheral objects, focus on several objects at the same time, or fixate solely one object. The presented toolkit incorporated three main design characteristics that can be manipulated: abstraction to ease the consumption of information, notification levels to differentiate importance of information, and transitions as effect to attract attention. Dealing with or even demanding user attention these displays can be classified as attention-aware.

According to Vertegaal (2003) such attentive user interfaces are distinguished by their focus on design for user attention and comprise means to measure, model, and interpret user attention. In contrast to other user interfaces they are capable to structure interaction accordingly, e.g. by first requesting

attention from the periphery and then wait for acknowledgement before capturing the focus of attention. The measured attention indicators are among others presence, proximity, orientation, activity, or gaze.

Starting from this interactional perspective (Michelis & Müller, 2011) derived their audience funnel framework after deploying and observing a public display installation. The framework is based on attention indicators and allows modelling the interaction between displays and their audience. The possible interaction is covered in six different phases. These phases are passing-by, viewing and reacting, subtle interaction, direct interaction, multiple interactions, and follow-up actions. The authors also discussed means to overcome the thresholds between the phases Müller, Alt, Michelis, & Schmidt (2010). Specifically they proposed to raise attention to reach the second phase, arouse curiosity to reach the third phase, and further motivate the audience to reach the other phases. Finally the authors suggested that public displays need to be balanced right to capture attention without annoying the audience and that the design should allow an effortless transition from one phase to the next in line with a shift from implicit to more explicit interaction modalities.

As an extension of the audience funnel framework (Wang, Boring, & Greenberg, 2012) presented the peddler framework. The framework does not only model distinct interaction states, but continuously captures and reacts upon the user's interest and attention state. Based on this information the framework adapts the display content and tries to attract, retain, and if necessary reacquire attention. Targeting on advertisers the framework rests upon the AIDA model (Strong, 1925) a basic strategy from advertising and marketing: attract attention, maintain interest, create desire, and lead customers to action. Consequently Wang et al. (2012) also implemented a prototype of a public advertising display based on their framework with the goal "to lead the [user] into a more attentive stage, ideally resulting in a purchase." To do so a product animation tries to attract the attention from persons passing by the display. The prototype captures persons in front of the display, including their position and orientation relative to the display. Whenever a person starts to look at the display the animation is slowed down. The appearance of personalised product recommendations then reflects the subtle interaction phase, followed by a direct interaction phase where the person is allowed to explore product details (and purchase the

product). The display also reacts when a person loses interest, either by starting to animate the product or showing different products. Discussing the prototype the authors highlighted once more the importance of choosing an appropriate design that allows drawing the user attention in a subtle and unobtrusive way, depending on the context of use. They concluded that for an effective attention-aware display design, the effort invested into shifting between the periphery and the focus of attention should not outweigh the importance of the presented information in relation to the primary task or activity.

1.2 Purpose

Following up previous research this empirical study was designed to further investigate the identified research challenges. The main objective was to find a realistic and reliable evaluation setting that allows controlling confounding variables and gain insights into the interplay between display design, user attention, and knowledge acquisition. The gathered results suggested designing a prototype, based on ambient display characteristics, that retains user interest over time. Positioned somewhere in between inattention, divided attention, and focused attention the prototype should use indexical representation to convey comprehensive and versatile information on a lower level of abstraction. Based on the existing taxonomies and frameworks, the prototype should furthermore implement interruptive notifications with abrupt transitions to attract attention and gradual transitions to retain attention. The assumption was that such an attention-aware design amplifies the user attention towards the display and thus facilitates the knowledge acquisition. The following research questions and hypotheses were derived and tested with the implemented prototype:

- (1) Does an attention-aware display design capture the user's focus of attention earlier and longer?

Hypothesis #1. The user attention is attracted earlier and retained significantly longer when applying an attention-aware display design. Thus the shift to the focus of attention is earlier and the time in the focus of attention is larger within the group exposed to the attention-aware display design than in the control group.

- (2) Does an attention-aware display design influence the knowledge gain?

Hypothesis #2a. The user attention towards the display is correlated to the possible knowledge gain. Thus the amount of knowledge gained by a user increases with the amount of attention paid towards the display.

Hypothesis #2b. The attention-aware design of a display facilitates the knowledge gain. Thus the amount of knowledge gained by a user is larger within the group exposed to the attention-aware display design than in the control group.

(3) Does an attention-aware design meet the general ambient display requirements?

Hypothesis #3. The attention-aware design of a display does not influence the display's unobtrusiveness and glanceable information presentation. Thus there is no significant difference in the rated disruptiveness, the required mental effort for filling in a questionnaire and for paying attention to the display, as well as the number of shifts between the periphery and the focus of attention between the group exposed to the attention-aware display design and the control group.

2. Method

For the experimental variation the display's ability to attract and retain user attention was defined as independent variable. The variable could take one of two distinct states, either the display had the ability or not. The study implemented a posttest-only randomised experimental design with two groups, i.e. a treatment group exposed to an attention-aware display design and a control group. The dependent variables measured were knowledge gain and user attention. Both variables were measured after the experiment without an initial pretest measurement. Thus both variables were depicted by individual values relative to all other participants.

2.1 Materials

For the experiment a display prototype that corresponded to the main characteristics of ambient displays was used. The display was designed to deliver information in the periphery of attention, while still being able to move between the periphery and the focus of attention. Each display consisted of a 27-inch Apple

iMac desktop computer with built-in speakers and webcam but without attached keyboard or mouse. The speakers were used to send out audio notifications, while the webcam was used to enhance the functionality of the computer with a custom-built attention sensor. The sensor measured (during the experiment) whether a participant looked at and thus attended the display. This concept is based on the work of Shen, Eades, and Moere (2007) but differs in its technical implementation. Shen et al. used the Intel OpenCV face detection/recognition programme in combination with an IR sensor. They also evaluated the accuracy of their implementation (Shen et al., 2008) and calculated an error rate of around 30%. The sensor used in the context of this study was built using the Processing¹ development environment and the open source computer vision library OpenCV for Processing. Using the included standard image processing capabilities of Processing the sensor captured webcam images and applied the OpenCV frontal face detection on each image to detect people facing the display. Finally the sensor's functionality was further enhanced to make the display attention-aware.

The display presented a slideshow with information and guidelines for first responders in emergency situations, especially in cases of cardiac arrest. The slideshow was compiled based on information developed in the context of a related project (Kalz et al., 2013) and consisted of 12 different slides. The slides depicted generic information about the project, the causes and symptoms of cardiac arrest, statistical probabilities, as well as detailed information on the recommended procedure to follow when witnessing cardiac arrest. The presented information was solely visual and static not containing any additional audio or video. **Error! Reference source not found.** shows the experimental setting with the display prototype. The setting was designed to reproduce a waiting room. Apart from the display the setting included a table with chairs, plants, books on the table and in a bookshelf, and pictures.

2.2 Instruments

A questionnaire was used to measure several individual components directly after the experiment. The questionnaire took about 10 minutes to complete and was not anonymous which allowed relating the results to the observational data gathered in the context of the study. The questionnaire was constructed

¹ <http://processing.org>

specifically for this study and contained demographic items (i.e. age and gender), items related to the individual perception and comprehension of the display, as well as items determining the mental effort (i.e. the amount of working memory capacity) invested by the participants in the course of the study. Most of these items used rating scales. To capture the participants' perception and comprehension several 5-point scaled items provided an open range of choices from 1 (not at all) to 5 (completely). The selection of items was based on an evaluation framework of ambient displays introduced by Holmquist (2004). The mental effort was measured as suggested by Paas (1992) using one 9-point scaled item providing an open range of choices from 1 (very, very low mental effort) to 9 (very, very high mental effort). This scale proved to be a reliable measure for the effort invested by an individual to handle a specific task and reflects the actual cognitive load (Paas, Tuovinen, Tabbers, & Van Gerven, 2003). Requiring minimal cognitive load is one of the core requirements of ambient displays and also one of the heuristics used when evaluating their usability and effectiveness (Mankoff et al., 2003). At the beginning of the questionnaire the participants were asked to rate their investment during the experiment, e.g. when paying attention to the information display. Furthermore the questionnaire contained test items to determine an individual knowledge score. These items consisted of single-choice, multiple-choice, or open questions. The given answers were then encoded, scored independently if correct, and aggregated to form a final score. Additionally participants were asked to indicate their level of prior knowledge and personal experience either as none, novice, or advanced. At the end of the questionnaire the participants were asked once more to rate their mental effort investment when performing this knowledge test. The different types of questionnaire items were not mixed when forming components to allow a consistent analysis. **Error! Reference source not found.** lists the components, some sample questions used in the questionnaire, and their type.

The attention paid towards the display was measured per participant using the custom-built attention sensor. The sensor calculated two related attention measures: a time-based linear measure depicting the amount of attention paid and a proportional measure depicting the level of attention reached. Without noticing, participants increased their amount of attention towards the display (maximum 5 points per

second up to a total of 3000 points). The level of attention increased with an exponentially increase of these attention points. Thus when reaching pre-defined thresholds in total 12 consecutive attention levels could be reached. The threshold y to unlock the next level was calculated using the exponential function $y_{\text{currentLevel}} = 1.75^{(\text{currentLevel}-1)} * 10$. The function was inspired by existing well elaborated game mechanisms and then developed on top of the specifications given by the custom-built attention sensor. Every participant started on the first level of attention. To reach the second level a participant needed $y_2 = 1.75^1 * 10 = 17.5$ attention points. Thus the levelling process started after a participant paid attention for around 3-4 seconds. The relative time needed to reach the second level equals the individual moment in time when the display first moved into the participant's focus of attention. To reach the last level a participant then needed at least $y_{11} = 1.75^{10} * 10 = 2694$ attention points. The results were automatically logged with additional information in two separate log files per participant. An event log file stored occurring events during the experiment, such as start of the experiment, given notification, attention towards the display, reaching a level of attention, and end of the experiment. Thereby each entry listed the event name, the experimental group, the current amount of attention paid, the achieved attention level, and the event time relative to the start of the experiment. Furthermore a session log file stored aggregated information at the end of each experimental session. Thereby each entry listed a session identifier, the experimental group, the starting time, the end time, the total amount of attention paid, the achieved attention level, the received notifications, an accumulated count depicting when the display moved in the focus of attention, and an accumulated count depicting when the display moved back to the periphery of attention. Several parameters could be derived for analysis from the gathered log data:

- Total amount of attention paid and thus the amount of time spent in the focus of attention.
- Achieved levels of attention and the relative time when achieving the levels.
- Number of shifts between the periphery and the focus of attention.

In addition to the measurement instruments, the experimental session was also observed from an adjacent room that allowed observing the participant and the display through a one-way glass window. The observations were recorded in a written protocol for each participant. In doing so the participants' main

and side activities during the experiment, such as reading a book or using a mobile phone, were captured. The protocol also captured participants' reactions on the experimental variation as well as possible malfunctions of the custom-built attention sensor and the display.

2.3 Participants

The target audience for this study is the general public. For practical reasons the actual experiment was conducted at a university campus. All employees working at the campus were asked to participate. In total 52 employees (20 female, 32 male; $M = 44.62$ years, $SD = 10.94$) agreed to participate. The participants were distributed to either the treatment or the control group at random. The groups consisted of 26 participants each. None of the participants had a medical or health-related background. Due to the assignment procedure the study implemented an experimental research design (Campbell, Stanley, & Gage, 1963). Participation in the experiment was voluntary. Informed consent was obtained, i.e. participants were informed about the general purpose of the experiment. Due to the ambient nature of the display the participants were not asked to watch out for any specific treatment. Directly after the experiment the participants were debriefed and the situation was resolved.

2.4 Procedure

The experiment was performed in individual sessions with single participants. The participants were welcomed and informed that they would participate in an experimental study evaluating a developed toolset. Furthermore they were told that setting up the toolset for evaluation would take some minutes, which should be used to already fill in a pre-questionnaire for the upcoming evaluation. They were then accompanied to the prepared waiting room, asked to have a seat, requested to fill in the pre-questionnaire, and then left alone. After that the information slideshow was started remotely. Each slide was presented for 25 seconds to allow the participants to capture all presented information. The complete slideshow was presented twice. Thus the experimental session lasted 600 seconds (= 10 minutes) in total. This represented a reasonable waiting time for the participants.

The given pre-questionnaire was used to provide the participants with a simple primary task to focus on during the experiment. The items were not related to the experiment and contained generic

demographic questions as well as an open-ended question to elaborate on. The pre-questionnaire took about 3-5 minutes to complete.

Depending on the assigned group the display reacted differently during the slideshow. The participants within the control group were exposed to a display just presenting the slides. The participants within the treatment group were exposed to the same with the variation that the display additionally tried to attract and retain their attention. The custom-built attention sensor was used to determine if the display was in the participant's periphery or focus of attention. When in the periphery the display tried to attract attention by sending audio notifications and showing a visual feedback. For the audio notification a short sample of a male voice saying 'Psst!' was looped in constant intervals of 12.5 seconds. As visual feedback a rectangular shape covered the lower area of the slides without hiding any content. The shape was filled with colour generating a blinking effect by randomly changing colour maximum several times per second. **Error! Reference source not found.**^a shows a sample slide with the overlaid visual feedback. Both the audio notification and the visual feedback vanished when the display moved to the participant's focus of attention. Instead, the display then tried to retain attention by providing visual and auditory feedback using gamification. When reaching the 12 consecutive levels of attention respective achievements could be unlocked. Rectangular shapes (red coloured in compliance to the slide design) illustrated these achievements. The shapes also covered only the lower area of the slides without hiding any content. Additionally, when unlocking an achievement a short audio confirmation was played as auditory feedback. For the audio confirmation a short sample consisting of a simple melodic line with 3 tones was used. **Error! Reference source not found.**^b shows a sample slide with overlaid achievements. The unlocked achievements vanished when the display moved back to the participant's periphery of attention.

The main game design patterns (Björk & Peitz, 2007) adapted for this gamification were rewards, experimenting, predictable consequences, outcome and progress indicators, as well as perceived chance to succeed. By providing positively received rewards for paying attention (unlocking achievements), the participants should be enabled to finally predict the consequences of their actions after learning about the

causes and effects of their performed actions. The vanishing visual feedback indicated the outcome when paying attention. The illustrated achievements then indicated the current progress towards an undefined goal and thus the perceived chance to succeed when paying more attention.

After the experimental session the participants were visited and taken to another room for further evaluation. They were asked if they had noticed something in the room, made aware about the experiment, guided through the following questionnaire, and finally informed about privacy and confidentiality regarding the gathered questionnaire and attention data.

3. Results

The assumptions of parametric data were checked prior to analysis. The data was appropriate as the dependent variables were independently measured at interval level and the used factor was categorical. The distribution of the sample was tested with Kolmogorov-Smirnov tests. The homogeneities of variance were tested with Levene's tests. Independent *t*-Tests as well as analyses of covariance (ANCOVA) were conducted to compare the mean values of the two groups and determine whether the differences are statistically significant. These parametric tests were considered as robust procedure producing reliable statistics when group sizes are equal. Thus the data violating the assumptions was not transformed. The effect sizes were estimated using Pearson's correlation coefficient (*r*) and partial eta squared (η^2).

All participants noticed the display in the room and were aware what kind of information was presented ($M_{\text{Total}} = 4.13$, $SD = 0.95$). They did understand the presented information ($M_{\text{Total}} = 3.96$, $SD = 1.03$), considered the used presentation as appealing ($M_{\text{Total}} = 3.40$, $SD = 1.11$), and were satisfied with the amount and granularity of the information ($M_{\text{Total}} = 3.42$, $SD = 1.11$). Furthermore they rated the information moderately useful and relevant ($M_{\text{Total}} = 2.98$, $SD = 1.41$). Independent *t*-Tests were conducted. All variances were equal. The tests revealed a significant difference between the groups on the rated satisfaction with the amount and granularity of the information, $t(50) = -2.06$, $p = .044$, $r = .28$. On average the group exposed to the attention-aware display design rated satisfaction higher ($M_{\text{Treatment}} = 3.73$, $SD = 0.87$) than the control group ($M_{\text{Control}} = 3.12$, $SD = 1.24$).

The first hypothesis was related to the user attention towards the display, determined by the relative time when participants reached the second level of attention and the total amount of attention paid. In total the average time for reaching the second attention level was $M_{\text{Total}} = 222.38$ seconds ($SD = 151.28$). Two participants (one from each group) failed to reach the second level and were treated as missing values. On average, the group exposed to the attention-aware display design reached the level significantly earlier ($M_{\text{Treatment}} = 155.63$ seconds, $SD = 118.63$) than the control group ($M_{\text{Control}} = 289.12$ seconds, $SD = 152.87$), $t(48) = 3.45$, $p = .001$, $r = .45$. The treatment group also reached all other levels earlier. The effect is significant up until level 8. No participant reached level 12. No participant in the control group reached level 11. Missing values of the time for reaching a level were excluded. The times for reaching the levels were normally distributed within the control group. The variances of the time for reaching the levels were equal. Table 2 lists the descriptive group statistics for all levels of attention reached and the respective t -Test results.

In total the average amount of attention paid was $M_{\text{Total}} = 502.06$ ($SD = 486.55$). The group exposed to the attention-aware display design paid a significantly larger amount of attention ($M_{\text{Treatment}} = 709.31$, $SD = 534.18$) than the control group ($M_{\text{Control}} = 294.81$, $SD = 328.98$), $t(50) = -3.37$, $p = .002$, $r = .43$. The amount of attention was normally distributed within the treatment group. The clustered bar chart in **Error! Reference source not found.** illustrates this observation by showing the groups' mean attention per slide.

The second hypothesis was related to the knowledge gain, determined by individual knowledge test scores. First, a correlation analysis was conducted to examine the relationship between attention and knowledge gain. The knowledge test score was significantly correlated with the amount of attention paid towards the display, $r = .54$, $p < .001$. Pearson's correlation coefficient was used as the variables were measured at interval level. The scatterplot in **Error! Reference source not found.** illustrates the relationship. The plot depicts the participants' attention against their knowledge test scores. The regression line summarises the relationship. The more attention a participant paid towards the display, the higher the participant's score in the knowledge test and thus the amount of knowledge gained. Calculating

the coefficient of determination by squaring the correlation coefficient reveals that user attention might be accounted for 29% of the variance in knowledge test score.

Second, the influence of the attention-aware display design on the knowledge gain was analysed. In total the average knowledge test score was $M_{\text{Total}} = 25.90$ ($SD = 6.84$). The group exposed to the attention-aware display design scored higher ($M_{\text{Treatment}} = 27.69$ seconds, $SD = 6.36$) than the control group ($M_{\text{Control}} = 24.12$ seconds, $SD = 6.95$). The boxplot in **Error! Reference source not found.** illustrates the group differences. The plot shows the median, the interquartile range, as well as the upper and lower quartile scores. A one-way ANCOVA was conducted. The knowledge test score was normally distributed within each group and the variances were equal. The indicated prior knowledge level was included in the analysis as confounding variable (covariate). It was appropriate to include this covariate, as the treatment and control group means of the variable are not significantly different. Furthermore the assumption of homogeneity of regression slopes was met, as there was no significant interaction effect between the factor and the covariate. The analysis revealed that the covariate, prior knowledge level, was significantly related to the knowledge test score, $F(1, 49) = 9.79$, $p = .003$, $r = .41$. There was also a significant effect of the attention-aware display design on the average knowledge test score after controlling for the effect of the prior knowledge level, $F(1, 49) = 4.41$, $p = .041$, $\text{partial } \eta^2 = .08$.

The third hypothesis was related to the requirements met by the display design, determined by the perceived disruptiveness, the rated mental effort required for filling in a questionnaire and for paying attention to the display, as well as the number of shifts between the periphery and the focus of attention. In total the average perceived disruptiveness was $M_{\text{Total}} = 2.29$ ($SD = 1.32$). The group exposed to the attention-aware display design perceived the display more disruptive ($M_{\text{Treatment}} = 3.08$, $SD = 1.20$) than the control group ($M_{\text{Control}} = 1.50$, $SD = 0.91$). In total the average rated mental effort required for filling in a questionnaire was $M_{\text{Total}} = 3.60$ ($SD = 1.71$). The group exposed to the attention-aware display design rated the effort lower ($M_{\text{Treatment}} = 3.46$, $SD = 1.68$) than the control group ($M_{\text{Control}} = 3.73$, $SD = 1.76$). In total the average rated mental effort required for paying attention to the display was $M_{\text{Total}} = 3.63$ ($SD = 1.89$). The group exposed to the attention-aware display design rated the effort higher ($M_{\text{Treatment}} = 4.12$,

$SD = 1.97$) than the control group ($M_{\text{Control}} = 3.15$, $SD = 1.71$). In total the average number of shifts between the periphery and the focus of attention was $M_{\text{Total}} = 10.42$ ($SD = 8.00$). The group exposed to the attention-aware display design shifted more often ($M_{\text{Treatment}} = 12.58$, $SD = 8.58$) than the control group ($M_{\text{Control}} = 8.27$, $SD = 6.88$). Independent t -Tests were conducted. The variables measured with scale items were also considered as interval data. The time for unlocking was normally distributed within the treatment group. The rated mental effort for paying attention to the display was normally distributed within the treatment group. The variances of both rated mental efforts and the number of shifts between the periphery and the focus of attention were equal. The tests revealed a significant difference between the groups on the perceived disruptiveness, $t(50) = -5.36$, $p < .001$, $r = .60$. There was no significant effect of the attention-aware display design on the rated mental effort required for filling in a questionnaire and for paying attention to the display nor for the number of shifts between the periphery and the focus of attention.

4. Discussion

The first research question was whether an attention-aware display design could capture the user's focus of attention. The corresponding hypothesis stated that such a design attracts the attention earlier and retains the attention longer. The results show clear evidence to support this hypothesis. The group exposed to the attention-aware display design reached the second attention level significantly earlier than the control group, on average in almost half of the time. When looking at the slideshow timing, the participants in the treatment group were already attracted to pay attention after 2-3 minutes halfway through the first set of slides, while the participants in the control group on average started to pay attention after 5 minutes and thus not before the end of the first set of slides. Here the used audio notification and gamification came into effect. Most of the participants started filling in the given pre-questionnaire and focused on that. When the participants in the treatment group received the first notification the majority intuitively looked for the source and found the display. The visual feedback then quickly confirmed this guess and the achievement unlocking further rewarded the action. The participants

in the treatment group also reached all other levels of attention (up to level 8) significantly earlier. For level 9 and 10 the group sizes get very small and the differences are not significant anymore. The number of participants reaching the levels decreased constantly from the third attention level on with a greater decrease in the control group. Only three participants from the treatment group finally reached level 11. The reported effect size increases from medium to high with each level and then decreases again for level 8. This development can be partly explained by the occurring novelty effect (Clark & Sugrue, 1988) when introducing the different elements to the treatment group. While experimenting around with the design the participants paid more attention. Consequently, the attention was also retained longer in the treatment group. The participants paid significantly more attention than the control group, on average more than twice as much. In relation to the possible maximum, the attention-aware display design captured almost 25% of the participants' attention compared to only 10% in the control group.

A limitation of the study is related to the question whether the observed effect is sustainable when participants are exposed to the same attention-aware display design multiple times. This has not been examined in this study but should be considered for future research. Another limitation of the presented study is related to the attention measurement. The user attention is measured conceptually assuming that a user pays attention to the display and the information presented whenever looking at it. Using the custom-built attention sensor this process is even more abstracted to a level where facing the display frontally equals looking at the display and thus paying attention. Although this broad abstraction is confirmed by previous and related research, there are much more reliable measures available. A solution could be eye tracking, a technique to measure eye movements and relate them to an area on a screen or display (Holmqvist et al., 2011). The collected movement indicators, such as fixation count and total fixation duration, can be abstracted to precisely allocate user attention, possibly also for multi-user scenarios. Just recently, Yang, Chang, Chien, Chien, & Tseng (2013) investigated the visual attention of students during a multimedia presentation in a classroom. Even though eye tracking becomes less intrusive with latest technical developments (e.g., mountable remote trackers, tracking glasses), it is still more intrusive compared to the sensor method used in this study.

After verifying that an attention-aware display design captures the user's focus of attention, the next research question asked whether this would also have an impact on the user's knowledge gain. The hypotheses were that there is a correlation between user attention and knowledge gain in general and that an attention-aware display design can facilitate this gain. The results show evidence to support both hypotheses. There is a positive relation between the amount of attention a participant paid towards the display and the participant's score in the knowledge test. Overall when participants paid more attention, their knowledge test scores (and thus knowledge gains) were higher. User attention might be accounted for only 29% of the variance in the knowledge score, so there are confounding variables that need to be considered, such as the prior knowledge level. Even taking this into account the group exposed to the attention-aware display design still scored significantly higher than the control group. This effect is medium in size. In conclusion, an attention-aware designed ambient learning display using indexical representation and interruptive notifications can attract and retain attention in such a way that the acquisition of knowledge (i.e. the comprehension of the presented information) is effectively facilitated.

The study has some limitations in this context. The presented information was mainly limited to factual knowledge. This could be improved by addressing also the conceptual or procedural knowledge dimensions and/or more complex cognitive process dimensions as defined by Anderson & Krathwohl (2001). Finally the evaluation does not go beyond the simple recall of acquired information and thus does not reveal insights into long-term learning effects, e.g. regarding a possible cognitive transfer of knowledge.

The last research question then asked whether such an attention-aware display design still meets the general ambient display requirements, such as to be unobtrusive and glanceable. The hypothesis stated that there would be no difference caused by the design. The results only partly show evidence to support this. The group exposed to the attention-aware design rated the display significantly more distracting than the control group. Here again the used audio notification and gamification came into effect. What proved to be effective to attract and retain user attention on the one hand is perceived too disruptive on the other. An adjustment is needed to find the right balance. When it comes to the perceived mental effort required

for filling in the questionnaire while paying attention to the display no significant difference between the groups could be found. Although this lack of significance supports the hypothesis, the actual group ratings are controversial. For filling in the questionnaire the treatment group rated the mental effort required lower than the control group. For paying attention to the display this turned as the treatment group rated the mental effort required much higher than the control group. This gives an indication regarding the fine interplay between primary tasks or activities and different forms of attention (inattention, divided attention, focused attention). A higher attention demand lowers the effort invested in the primary task and vice versa. The hypothesis is further supported when examining the number of shifts between the periphery and the focus of attention. Due to the notifications the treatment group shifts more but the difference between the groups is not significant.

This is related to another limitation of the study. It can be argued that the used interruptive notification design was too obtrusive (even distracting) and that the observed effects are merely a result of this. The experimental prototype was design in such a way that the variations depicted borderline manifestations of the manipulated ambient display characteristics. Reproducing the effects future designs should gradually decrease the level and frequency of notification to find a right balance.

5. Conclusions

This paper described a study to evaluate the effect of an attention-aware ambient display design in a learning context. As the second empirical study into the research and development of ambient learning displays, this study was designed to further investigate previously identified research challenges. The main purpose was to design a noticeable, unobtrusive, and comprehensive display that is capable to retain user interest over time, evaluate this design in an authentic setting controlling confounding variables, and gain insights into the interplay between display design, user attention, and knowledge acquisition. The results provide evidence that an attention-aware display design attracts and retains user attention more effectively and that there is a positive relation between knowledge gain and user attention. Furthermore the design significantly facilitated the acquisition of knowledge. The experimental setting proved to be a

good compromise between internal validity and ecological validity. The method holds the potential to be used for future evaluation studies of public displays without violating ethical guidelines. For the efficient increase of basic knowledge about activities of first responders in case of cardiac arrest the introduced ambient learning displays can be applied in semi-public contexts like waiting rooms.

The study continues the work on ambient learning displays and makes an important contribution to the field of technology-enhanced learning. We see the approach presented as a promising way to align the specific characteristics of ambient displays with the contrasting demands on technology in a learning context. Considerably more work will need to be done to determine the cognitive transfer of knowledge on the long term. Furthermore the sustainability of the observed attention effect needs to be investigated, possibly using the same experimental set up for repeated measures just altering the presented information. The next step regarding the research and development of ambient learning displays is then evaluating the next phases of audience interaction, i.e. direct and multiple interactions. Especially the effects of implicit and explicit interaction modalities should be explored in further research.

References

- Alt, F., Schneegaß, S., Schmidt, A., Müller, J., & Memarovic, N. (2012). How to evaluate public displays. In *Proceedings of the 2012 International Symposium on Pervasive Displays - PerDis '12* (pp. 1–6). New York, NY: ACM Press.
- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's Taxonomy of educational objectives*. New York, NY: Longman.
- Björk, S., & Peitz, J. (2007). Understanding Pervasive Games through Gameplay Design Patterns. In *DiGRA 2007 - Situated Play*. In *Proceedings of the 3rd International Conference of the Digital Games Research Association* (pp. 440–448). The University of Tokyo.
- Börner, D., Kalz, M., & Specht, M. (2011). Thinking outside the box - A vision on ambient learning displays. *International Journal of Technology Enhanced Learning*, 3(6), 627–642.
- Börner, D., Kalz, M., & Specht, M. (2013a). Beyond the channel: A literature review on ambient displays for learning. *Computers & Education*, 60(1), 426–435.
- Börner, D., Kalz, M., & Specht, M. (2013b). It Doesn't Matter, But: Examining the impact of ambient learning displays on energy consumption and conservation at the workplace. *Manuscript Accepted for Publication*.

- Campbell, D. T., Stanley, J. C., & Gage, N. L. (1963). *Experimental and quasi-experimental designs for research* (p. 84). Boston, MA: Houghton, Mifflin and Company. Retrieved from http://www.uky.edu/~clthyn2/PS671/CS_part1.pdf
- Clark, R. E., & Sugrue, B. M. (1988). Research on instructional media, 1978-1988. In D. Ely (Ed.), *Educational Media Yearbook 1987-88*. Littleton, CO: Libraries Unlimited.
- Holmqvist, L. E. (2004). Evaluating the comprehension of ambient displays. In *CHI '04 extended abstracts on Human factors in computing systems* (pp. 1545–1545). New York, NY: ACM Press.
- Holmqvist, K., Nyström, N., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking: a comprehensive guide to methods and measures*. Oxford: Oxford University Press.
- Kaikkonen, A., Kekäläinen, A., Cankar, M., Kallio, T., & Kankainen, A. (2005). Usability testing of mobile applications: a comparison between laboratory and field testing. *Journal of Usability Studies*, 1(1), 4–16. Retrieved from <http://lib.tkk.fi/Diss/2009/isbn9789522481900/article6.pdf>
- Kalz, M., Klerx, J., Parra, G., Elsner, J., Haberstroh, M., Ternier, S., ... Specht, M. (2013). EMuRgency: Addressing cardiac arrest with socio-technical innovation in a smart learning region. *Interaction Design and Architectures Journal*, 17(2), 77–91. Retrieved from http://www.mifav.uniroma2.it/inevent/events/idea2010/index.php?s=10&a=11&link=ToC_17_P
- Mankoff, J., Dey, A. K., Hsieh, G., Kientz, J., Lederer, S., & Ames, M. (2003). Heuristic evaluation of ambient displays. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 169–176). New York, NY: ACM Press.
- Matthews, T., Dey, A. K., Mankoff, J., Carter, S., & Rattenbury, T. (2004). A toolkit for managing user attention in peripheral displays. In *Proceedings of the 17th annual ACM symposium on User interface software and technology - UIST '04* (Vol. 6, pp. 247–256). New York, NY: ACM Press.
- Michelis, D., & Müller, J. (2011). The Audience Funnel: Observations of Gesture Based Interaction With Multiple Large Displays in a City Center. *International Journal of Human-Computer Interaction*, 27(6), 562–579.
- Müller, J., Alt, F., Michelis, D., & Schmidt, A. (2010). Requirements and design space for interactive public displays. In *Proceedings of the international conference on Multimedia - MM '10* (p. 1285). New York, NY: ACM Press.
- Paas, F. G. W. C. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive load approach. *Journal of Educational Psychology*, 84, 429–434.
- Paas, F. G. W. C., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, 38, 63–71.
- Pousman, Z., & Stasko, J. (2006). A taxonomy of ambient information systems: four patterns of design. In *Proceedings of the working conference on Advanced visual interfaces* (pp. 67–74). New York, NY: ACM Press.

- Shen, X., Eades, P., & Moere, A. Vande. (2007). Intrusive and Non-intrusive Evaluation of Ambient Displays. In *Proceedings of the 1st International Workshop on Ambient Information Systems*.
- Shen, X., Moere, A. Vande, Eades, P., & Hong, S. (2008). The long-term evaluation of Fisherman in a partial-attention environment. In *Proceedings of the 2008 conference on BEyond time and errors: novel evaluation methods for Information Visualization*. New York, NY: ACM Press.
- Strong, E. (1925). *The psychology of selling and advertising*. New York, NY: McGraw-Hill. Retrieved from http://scholar.google.de/scholar?q=author:strong+selling+advertising&btnG=&hl=de&as_sdt=0,5#0
- Vertegaal, R. (2003). Attentive User Interfaces. *Communications of the ACM*, 46(3), 31–33.
- Wang, M., Boring, S., & Greenberg, S. (2012). Proxemic peddler: A Public Advertising Display that Captures and Preserves the Attention of a Passerby. In *Proceedings of the 2012 International Symposium on Pervasive Displays - PerDis '12* (pp. 1–6). New York, NY: ACM Press.
- Werbach, K., & Hunter, D. (2012). *For The Win: How Game Thinking Can Revolutionize Your Business* (p. 148). Wharton Digital Press.
- Wisneski, C., Ishii, H., Dahley, A., Gorbet, M., Brave, S., Ullmer, B., & Yarin, P. (1998). Ambient displays: Turning architectural space into an interface between people and digital information. In *Proceedings of the First International Workshop on Cooperative Buildings* (Vol. 1370, pp. 22–32). New York, NY: Springer. Retrieved from <http://www.springerlink.com/index/744438x187j43417.pdf>
- Yang, F.-Y., Chang, C.-Y., Chien, W.-R., Chien, Y.-T., & Tseng, Y.-H. (2013). Tracking learners' visual attention during a multimedia presentation in a real classroom. *Computers & Education*, 62, 208–220.