Can computer vision cure display blindness? An investigation into the impact of tailoring advertisements to demographic attributes of passers-by

Saar Bossuyt
University College Leuven-Limburg
Floris De Feyter
KU Leuven
Kathleen Custers
University College Leuven Limburg

Acknowledgements:
We are thankful for Hilde Walravens, who helped to carry out both studies.

Cite as:
Bossuyt Saar, De Feyter Floris, Custers Kathleen (2021), Can computer vision cure display blindness? An investigation into the impact of tailoring advertisements to demographic attributes of passers-by. *Proceedings of the European Marketing Academy*, 50th, (93471)

Paper from the 50th Annual EMAC Conference, Madrid, May 25-28, 2021

EMAC
EUROPEAN MARKETING ACADEMY

Can computer vision cure display blindness? An investigation into the impact of tailoring advertisements to demographic attributes of passers-by

Display blindness, or the lack of attention for digital advertising displays due to low relevance expectations, is a major issue in the growing digital signage market. In this paper, we explore if recent developments in computer vision algorithms that allow to detect and classify demographic attributes of people passing by digital displays increase the relevance and hence the impact of digital signage. We set up two mobile eye-tracking studies to examine if digital advertisements tailored to the passerby's gender (Study 1; N=117) or age (Study 2; N=181) are considered more relevant, receive more visual attention, and are better recalled. Both studies yielded similar results: while advertisements congruent with the participants' demographic attributes were considered significantly more relevant than incongruent advertisements, this increased relevance was not translated into considerably more visual attention or better recall.

Keywords: *advertising effectiveness, digital signage, computer vision*
Track: *Advertising & Marketing Communications*

1. Introduction

Digital signage or Digital Out Of Home Advertising (DOOHA) are umbrella terms for digital displays in public spaces used for advertising or information purposes which can be adapted in real time (Bauer & Lasinger, 2014). The value of the digital signage market is growing, and is estimated to be worth about 14 billion US dollars in 2021 (Statista, 2020). In stark contrast to this growing trend, there is a huge lack of knowledge concerning the effectiveness of digital signage (Andrew et al., 2019; Cho, 2019). The call for more research is amplified by worrying results from previous studies on the effectiveness of digital displays that found evidence for a phenomenon called 'display blindness' (Ervasti et al., 2015; Huang et al., 2008; Müller et al., 2009). Display blindness implies that people do not pay attention to the content of digital displays, because they do not expect this content to be relevant to them (Ervasti et al., 2015; Müller et al., 2009; Ravnik & Solina, 2013). Finding a way to make the content of digital signage more relevant for the target audience is thus key.

Previous research describes two ways to make advertising more relevant: personalization and situationalization (Bauer & Lasinger, 2014). Personalization implies tailoring advertising based on demographic attributes from a (group of) consumer(s) (e.g., age, gender), while situationalization tailors advertising to situational characteristics (e.g. weather, location). Traditionally, situationalization was the preferred way to make digital signage more relevant (Bauer & Lasinger, 2014). However, recent developments in computer vision technology have turned personalization into a viable strategy as well. Advancements in deep learning have pushed many computer vision algorithms, like object detection, towards impressive performance in unconstrained situations (Ren et al., 2015). Moreover, the recent availability of large datasets consisting of pedestrian images with annotated demographic attributes like gender and age, allows us to develop deep learning-based algorithms that can classify such properties for a detected person (Deng et al., 2014).

In this paper, we examine if the impact of digital signage can be improved by employing a combination of deep learning-based person detection and person attribute classification to tailor advertisements to demographic characteristics of passers-by. In particular, we explore if such advertisements perform better in terms of relevance, visual attention, and recall.

2. Study 1

2.1 Method

The aim of the first study was to explore if adapting advertising content to a passerby's gender resulted in better advertising impact. We opted for gender as a characteristic as it is common to tailor advertisements for male or female versions of products (e.g. clothes or perfume) to the gender of the audience, and because it has been used in previous work on the impact of personalized advertisements as well (e.g., De Keyzer et al., 2015). We designed a 2 (Gender participant: male/female) x 2 (Gender ad: male/female) between-subjects experiment, implying that participants were either exposed to a clothing advertising that fitted with their gender (women: female clothes ad; men: male clothes ad) or that did not fit (vice versa) in a setting with high ecological validity. Participants were recruited to take part in a so-called navigation study on our campus (cover story: "Every year, many new students get lost on the campus. We want to solve this by investigating the gaze behavior of students navigating the campus"). After signing the informed consent, participants were fitted with eye-tracking glasses (Tobii Pro Glasses 2), and were instructed to walk to a room on the campus. Next to this room, we had previously installed a screen on which either the male or the female clothing ad could appear. After finishing the navigation exercise, participants were guided back to our stand to complete a follow-up survey.

Deep learning-based person detection and person attribute classification algorithms were used to determine which advertisement would appear on the screen. In particular, we created a computer vision pipeline that detected passers-by, classified the most salient detection, and decided based on this input which advertisement to display. To comply with the experimental design, we slightly adapted the algorithm to randomly show either the advertisement that matched the classified gender or the advertisement that did not match the classified gender. The pipeline ran on a laptop with an on-board GPU and a Logitech C310 webcam that provided the images, both connected to the screen. For person detection, we employed the YOLOv3 object detector (Redmon & Farhadi, 2018) due to its small inference time, which, evidently, was crucial in our real-time application. To ensure that the person detections only contained full-body detections, our pipeline ignored all person detections where the height was less than three times the width. Finally, the largest remaining person detection was passed onto the classification stage. The classifier consisted of a ResNet-50 convolutional neural network (He et al., 2016) trained on the PETA dataset (Deng et al., 2014). During training, we achieved 89.5% classification accuracy. Once the detection was classified, the classification result was appended to a circular buffer with a maximum capacity of five items, and each item expired after five seconds. The circular buffer was read every 10th of a second to decide the new state of the display (either a male or a female clothing ad). When the buffer was not entirely filled, the display remained in an ‘idle’ state, showing the campus logo.

We ran a pre-test to examine if gender congruent clothing ads (i.e. ads in line with one’s gender) were considered more relevant than gender incongruent clothing ads. Participants (n= 108; 54% male) saw both ads in a random order and were asked how relevant (1 ‘not relevant at all’ to 7 ‘very relevant’) they considered these. A 2 (Gender participant) x 2 (Gender ad) mixed ANOVA with relevance as dependent variable showed no significant main effects of Gender ad ($F(1,106) = 0.34$; $p = .564$) or Gender participant ($F(1,106) < 0.001$; $p = .986$). Only the Gender ad x Gender participant interaction was significant ($F(1,106) = 20.06$; $p < .001$): Both men ($\text{Men}_{\text{Congruent}}$: M=3.95, SD=1.55) and women ($\text{Women}_{\text{Congruent}}$: M=4.04, SD=1.78) found the gender congruent clothing ad more relevant than the gender incongruent clothing ad ($\text{Men}_{\text{Incongruent}}$: M=3.36, SD=1.59; $\text{Women}_{\text{Incongruent}}$: M=3.28, SD=1.54).

2.2 Results

In total, 172 students took part in this study. Unfortunately, several participants could not be included in the final sample. Early on in the experiment, we noticed that the screen was placed in a suboptimal position, making it not fully visible for participants taking a certain route to the classroom. We changed the screen’s position as of the 21st participant, but this implied that the first 20 participants had to be removed from the sample. Next, another 10 participants were removed because of technical issues with the eye-tracking equipment. Finally, the computer vision pipeline failed to classify 25 participants in time, causing the screen to remain idle instead of showing the male or female clothing advertisement. Hence, the final sample consisted of 117 participants (53% male).

Advertising impact was assessed with two commonly used metrics (Ciceri et al., 2020): visual attention (fixation count and duration or how often and long participants focused on the ad) and aided recall. We ran two 2 (Gender ad) x 2 (Gender participant) ANOVAs: one with fixation count and one with total fixation duration as dependent variable. There was no evidence that gender congruent ads received more visual attention than incongruent ads (see Table 1). In contrast with previous research stating that men have more attention for digital signage than women, we found that women (*M*=3.40; *SD*=3.45) fixated more often on both ads than men (*M*=2.08; *SD*=1.97) (Ervasti et al., 2015; Ravnik & Solina, 2013).

Table 1 - Fixation count and Total fixation duration 2x2 ANOVAs

Fixation count: Test of between subjects effects	*F*(1,113)	*p*
Gender ad (0 female; 1 male)	0.10	.747
Gender participant (0 female; 1 male)	6.73	.011
Gender ad x Gender participant	0.56	.454
Total fixation duration: Test of between subjects effects	*F*(1,113)	*p*
Gender ad (0 female; 1 male)	0.40	.528
Gender participant (0 female; 1 male)	2.91	.091
Gender ad x Gender participant	0.05	.833

For aided recall ("Have you seen a clothing advertisement during your walk?" No = 0; Yes = 1), we ran a binary logistic regression (see Table 2). The results were in line with the ones on visual attention: the gender of the advertisement and the gender of the participant did not significantly predict the likelihood of recalling the clothing advertisement.

Table 2 - Binary logistic regression Aided recall

				95% CI Odds Ratio		
	b(SE)	*Wald*(*df*)	*p*	Lower	OR	Upper
Constant	-0,31 (0.40)	0.61(1)	.435			
Gender ad (0 female; 1 male)	-0.83(0.59)	2.02(1)	.156	0.14	0.43	1.37
Gender participant (0 female; 1 male)	-0.88(0.59)	2.25(1)	.134	0.13	0.42	1.13
Gender ad x Gender participant	-0.75(0.85)	0.79(1)	.374	0.40	2.12	11.12
R^2= .03 (Cox-Snell), .04 (Nagelkerke); Model $\chi^2(3)$=3.59, *p*=.309						

3. Study 2

In Study 2, the research method was perfected in three ways. First, the conditions for viewing the advertisement were improved by designing an exercise in which participants had to wait near the screen rather than walk towards it. Previous research shows that people who are waiting tend to be more engaged with digital advertisements than walking people (Bauer et al., 2012). Second, we tried to increase the relevance of the ads by selecting advertisements related to an upcoming event on the campus and by tailoring them based on a more specific criterion than gender (profession: student or staff). Third, we manually changed the screen instead of working with the computer vision pipeline. We did so to avoid losing participants who were exposed to an idle screen instead of the ad, caused by the computer vision pipeline failing to classify them in time. Although this issue can be mended, we decided to focus first on finding more evidence for the effectiveness of the concept of 'advertisements tailored to demographic characteristics' before further tweaking the computer vision pipeline.

3.1 Method

To measure the effectiveness of ads fitted to the profession of passers-by, we designed a 2 (Advertised event: student/staff) x 2 (Profession participant: student/staff) between-subjects experiment, which implied that participants were either exposed to an advertisement for an event that fitted with their profession (student: student event; staff: staff event) or that did not fit (vice versa) in a setting with high ecological validity. Technically, an algorithm could detect the age of participants and hence classify their profession, making this a realistic concept test. The experiment took place in the coffee lounge of our campus, where we installed a screen right next to the coffee vending machine. All participants were students or staff members who were intercepted while waiting in line to buy coffee (cover story: "The owners of the coffee lounge would like to gain insights in the decision journey of their customers, making use of eye-tracking"). They signed an informed consent and were fitted with eye-tracking glasses (Tobii Pro Glasses 2). Next, they were instructed to buy coffee 'as usual', and subsequently return to the stand to complete a follow-up survey.

The advertisements shown on the screen were created by the event's organizers. As we wanted to avoid that our main sample contained both participants who were recently exposed to and questioned about the visuals in a pre-test and participants who were not, we decided to run a post-test (N=181; 43% male; 60% students) on relevance instead. A 2 (Profession participant) x 2 (Advertised event) ANOVA with relevance as dependent variable proved that our manipulation was successful: There was no significant main effect of Advertised event (F(1,177) = 2.33; p = .129), and the significant main effect of Profession participant (F(1,177) = 15.53; p < .001) was qualified by a significant Advertised event x Profession participant interaction effect (F(1,177) = 37.20; p < .001): Both students ($\text{Student}_{\text{Congruent}}$: M=3.98, SD=1.77) and staff ($\text{Staff}_{\text{Congruent}}$: M=5.44, SD=1.96) found the profession congruent event advertisement more relevant than the incongruent event advertisement ($\text{Student}_{\text{Incongruent}}$: M=2.76, SD=1.67; $\text{Staff}_{\text{Incongruent}}$: M=3.48, SD=1.66).

3.2 Results

In total, 182 people took part in our study. One person was excluded because of technical issues with the eye-tracking equipment. The final sample contained 181 participants (43% male; 60% students). As in Study 1, we measured advertising impact by visual attention and aided recall. We ran two 2 (Advertised event) x 2 (Profession participant) ANOVAs, one with fixation count and one with total fixation duration as dependent variable (see Table 3). The results were similar to those of Study 1: higher relevance was not translated into more visual attention. We did find that staff members fixated more often (M=10.59; SD=10.13) and longer (M=3.30; SD=3.53) on both ads than students (M=6.66; SD=8.59; M=2.21; SD=3.55).

Table 3 - Fixation count and Total fixation duration 2x2 ANOVAs

Fixation count: Test of between subjects effects	*F*(1,176)	*p*
Advertised event (0 staff; 1 students)	0.11	.738
Profession participant (0 staff; 1 students)	7.69	.006
Advertised event x Profession participant	0.17	.682
Total fixation duration: Test of between subjects effects	*F*(1,176)	*p*
Advertised event (0 staff; 1 students)	0.02	.900
Profession participant (0 staff; 1 students)	3.95	.048
Advertised event x Profession participant	0.60	.439

For aided recall (“Have you seen an advertisement for a student/staff event while buying coffee?” No = 0; Yes = 1) we ran a binary logistic regression (see Table 4). The results were similar to those on visual attention: advertising-profession congruency did not result in better recall, and staff members were more likely to recall either one of the ads than students.

Table 4 - Binary logistic regression Aided recall

				95% CI Odds Ratio		
	b(SE)	*Wald(df)*	*p*	Lower	OR	Upper
Constant	1.42 (0.42)	0.05(1)	.001			
Advertised event (0 staff; 1 students)	-0.13 (0.58)	4.65(1)	.818	0.28	0.88	2.73
Profession participant (0 staff; 1 students)	-1.10 (0.51)	0,05(1)	.031	0.12	0.33	0.91
Advertised event x Profession participant	0.16 (0.70)	11.39(1)	.820	0.30	1.17	4.62
R^2= .05 (Cox-Snell), .07 (Nagelkerke); Model $\chi^2(3)$=9.15, p=.027						

4. Discussion

The goal of this research was to find a solution for display blindness, or the fact that people do not pay attention to digital signage as they do not expect to see relevant content (Ervasti et al., 2015; Müller et al., 2009; Ravnik & Solina, 2013). In two studies, we explored if computer vision technology, which enables tailoring advertisements to demographic characteristics of passers-by (such as gender and age), could help to increase the relevance of the advertised content, which in turn could improve advertising effectiveness (measured in terms of visual attention and aided recall). In the first study, we found that even though both men and women found the gender congruent clothing ad more relevant than the incongruent clothing ad, they did not have more visual attention or better recall for the congruent than for the incongruent ad. In the second study, where viewing conditions for the advertised content were improved and the content was made even more relevant, the results were similar. Visual attention and aided recall were not better for congruent ads (student-student event; staff-staff event) than for the incongruent ads (student-staff event; staff-student event), even though congruent ads were considered more relevant than incongruent ads.

These findings pose an interesting challenge for future research on display blindness, as the question remains what, if not relevance, is (are) the true driver(s) for attention and recall of digital displays? In order to find the cure for display blindness, future studies should focus on this question first.

5. References

Andrew, H., Haines, H., & Seixas, S. (2019). Using neuroscience to understand the impact of premium digital out-of-home media. *International Journal of Market Research*, *61*(6), 588–600. https://doi.org/10.1177/1470785319851316

Bauer, C., Dohmen, P., & Strauss, C. (2012). A conceptual framework for backend services of contextual digital signage. *Journal of Service Science Research*, *4*(2), 271–297. https://doi.org/10.1007/s12927-012-0011-4

Bauer, C., & Lasinger, P. (2014). Adaptation strategies to increase advertisement effectiveness in digital media. *Management Review Quarterly*, *64*(2), 101–124. https://doi.org/10.1007/s11301-014-0101-0

Cho, J.-Y. (2019). A Study of the Effectiveness of Digital Signage: Importance of Customized Content. *Journal of Digital Convergence*, *17(6)*(6), 211–217.

Ciceri, A., Russo, V., Songa, G., Gabrielli, G., & Clement, J. (2020). A neuroscientific method for assessing effectiveness of digital vs. Print ads: Using biometric techniques to measure cross-media ad experience and recall. *Journal of Advertising Research*, *60*(1), 71–86. https://doi.org/10.2501/JAR-2019-015

De Keyzer, F., Dens, N., & De Pelsmacker, P. (2015). Is this for me? How Consumers Respond to Personalized Advertising on Social Network Sites. *Journal of Interactive Advertising*, *15*(2), 124–134. https://doi.org/10.1080/15252019.2015.1082450

Deng, Y., Luo, P., Loy, C. C., & Tang, X. (2014). Pedestrian attribute recognition at far distance. *MM 2014 - Proceedings of the 2014 ACM Conference on Multimedia*, 789–792. https://doi.org/10.1145/2647868.2654966

Ervasti, M., Häikiö, J., Isomursu, M., Isomursu, P., & Liuska, T. (2015). Digital Signage Effectiveness in Retail Stores. In J. Abascal, S. Barbosa, M. Fetter, T. Gross, P.

Palanque, & M. Winckler (Eds.), *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 9297, pp. 455–462). Springer International Publishing. https://doi.org/10.1007/978-3-319-22668-2_35

He, K., Zhang, X., Ren, S., & Sun, J. (2016). *Deep Residual Learning for Image Recognition*. http://image-net.org/challenges/LSVRC/2015/

Huang, E. M., Koster, A., & Borchers, J. (2008). Overcoming assumptions and uncovering practices: When does the public really look at public displays? *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 5013 LNCS*, 228–243. https://doi.org/10.1007/978-3-540-79576-6_14

Müller, J., Wilmsmann, D., Exeler, J., Buzeck, M., Schmidt, A., Jay, T., & Krüger, A. (2009). Display blindness: The effect of expectations on attention towards digital signage. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 5538 LNCS*(May), 1–8. https://doi.org/10.1007/978-3-642-01516-8_1

Ravnik, R., & Solina, F. (2013). Audience measurement of digital signage: Quantitative study in real-world environment using computer vision. *Interacting with Computers*, *25*(3), 218–228. https://doi.org/10.1093/iwc/iws023

Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. *ArXiv*. http://arxiv.org/abs/1804.02767

Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In *Advances in Neural Information Processing Systems* (Vol. 28).

Statista. (2020, February 26). *Value of the digital signage display market worldwide in 2016 and 2021 (in billion U.S. dollars)*. Https://Www.Statista.Com/Statistics/624103/Digital-Signage-Display-Market/.