WZB

Berlin Social Science Center

Christina Timko
Maja Adena

The Impact of Behavioral Design and Users' Choice on Smartphone App Usage and Willingness to Pay: A Framed Field Experiment*

Discussion Paper

SP II 2026–602

January 2026

*This is a revised version of the WZB Discussion Paper SP II 2023-302 "Transparent app design reduces excessive usage time and increases willingness to pay compared to common behavioral design: a framed field experiment", originally published by the same authors within the Research Unit Economics of Change.

Research Area
Markets and Choice

Research Group
Information, Incentives, Inequality

WZB Berlin Social Science Center
Reichpietschufer 50
10785 Berlin
Germany
www.wzb.eu

Christina Timko, Maja Adena
The Impact of Behavioral Design and Users' Choice on Smartphone App Usage and Willingness to Pay: A Framed Field Experiment
Discussion Paper SP II 2026–602
Wissenschaftszentrum Berlin für Sozialforschung (2026)

Affiliation of the authors:

Christina Timko
Ruhr-Universität Bochum

Maja Adena
Wissenschaftszentrum Berlin für Sozialforschung

Abstract

The Impact of Behavioral Design and Users' Choice on Smartphone App Usage and Willingness to Pay: A Framed Field Experiment

by Christina Timko and Maja Adena

Behavioral design in smartphone apps aims at inducing certain, monetizable behavior, mainly increased engagement, measurable by usage time. Such design is rarely transparent and often restricts users' ability to make alternative choices. In a framed field experiment, we document that behavioral design doubles app usage time compared to a version without behavioral elements. Providing users with choices—simply explained and conveniently adjustable design features—reduces usage time and increases their willingness to pay for the app. These findings suggest that offering choice could pave the way for new business models based on more responsible app design.

Keywords**:** *smartphone app, behavioral control, filtering algorithm, transparency and choice, self–determination, corporate social responsibility, field experiment.*

Keywords: smartphone app, behavioral control, filtering algorithm, transparency and choice, self–determination, corporate social responsibility, field experiment.

JEL classification: C93, O33, D83, L86, M14

---

* Christina Timko (dr.chris.timko@gmail.com) Ruhr-University Bochum, Universitätsstraße 150, 44801 Bochum, Germany, until 2023. Maja Adena (maja.adena@wzb.eu), WZB Berlin Social Science Center, Reichpietschufer 50, 10785 Berlin, Germany. CT thanks her partners in the GOAL project for their valuable comments (in alphabetical order): Yannik Borutta, Anja Folberth, Matthias Haag, Marc Hauer, Johannes Kevekordes, Carsten Orwat, Catharina Rudschies, Ingrid Schneider, and Katharina Zweig. CT especially thanks Kastanie Eins for the app development and Anne-Kathrine Kjær Christensen at Specifii in Denmark for her behavioral design workshop and her input on the user profile, concepts, and design ideas. The authors thank their students and testers for valuable feedback. Special thanks to Nicholas Ostrode and Malte Niederstadt for supporting the literature review, the preparation of the analysis work, and the conduction of the experiment and Katharina Dorn for supporting the analysis. CT thanks Michael Roos for entrusting her with this project. The data was collected via the MaXLab of the Otto von Guericke University in Magdeburg, Germany in November 2021. Data and programs are available at: [will be filled in upon acceptance]. The study was funded in line with the GOAL (Governance of and by Algorithms) project by the German Federal Ministry of Education and Research (BMBF) from 01/09/2019 to 31/12/2021, Funding ID: 01IS19020 (CT). Support by Deutsche Forschungsgemeinschaft through CRC TRR 190 (project number 280092119) and FOR 5675 (project number 518302089), and by Leibniz through P146/2022 and K595/2024 is gratefully acknowledged (MA).

1. INTRODUCTION

Smartphone apps have become a central part of daily life, generating significant revenue and consuming substantial user time. In 2023, over 6 million apps generated $550–900 billion USD in revenue, with an annual growth rate of 19.5%. Users spend an average of 4.8 hours daily on apps, primarily social and communication tools.[1] Most app business models achieve high conversion rates and viral growth relying on behavioral design (Alter, 2017; Lambrecht et al. 2014). Behavioral design is a design framework for programming (i.e., intentionally and systematically changing) human behavior by modifying the physical and digital environment (Combs & Brown, 2018). It can take various forms, including app appearance, usability, rewards, social features, progress monitoring, black-box algorithms, and personalization. It includes behavioral design elements from persuasive technology, nudging, behavioral interventions, and even aggressive and manipulative design (Michie et al., 2013; Oinas-Kukkonen & Harjumaa, 2008; Thaler, 2018).[2]

App providers commonly apply behavioral design to influence consumption decisions, nudge users to share personal data that can be monetized, or manipulate them into taking certain actions (Luguri & Strahilevitz, 2021). They personalize content and advertising to sell products (Boerman et al., 2017), make price discrimination possible (Acquisti et al., 2016), assist decision making (Kleinberg et al., 2018), nudge and influence users (Dellarocas, 2006), bind their attention and loyalty (Claussen et al., 2013), and create complex, much-tested tools with hidden business purposes (Montag et al., 2019). Putting user engagement but not user benefits at the heart of the design, behavioral design can exploit cognitive biases and induce harmful, risk-prone, or addictive behavior (Mosquera et al., 2020). It can induce users to share more data than they otherwise would

[1] https://web.archive.org/web/20241226101832/https://mindsea.com/app-stats/

[2] See also the overarching systematics that we created for this study in Table C1 in Appendix C

be willing to do (Acquisti et al., 2015). Moreover, one risk in digital media is digital addiction, where traditional metrics may overestimate consumer utility of social media apps (Allcott et al., 2020). Recent research provides evidence that social media are habit-forming and their excessive use is driven by self-control problems, with up to 31% of social media use attributable to such issues (Allcott et.al., 2022). Further studies find that news content personalization reduces knowledge (Beam, 2014) and contributes to the creation of echo chambers (Bail et al., 2018).

Behavioral design becomes problematic when it deliberately disregards user preferences, compromises freedom of choice, and undermines individual autonomy and the basic right of self-determination (Susser et al. 2018). In other words, if it lacks free and informed choice. Such designs are termed sludge, deceptive design, or dark patterns (Mathur et al. 2019; Thaler, 2018). Since users are not informed about, not asked to consent to, or unable to avoid such practices, they are addressed by new legislation that strengthens consumer protection, such as the newly enacted Digital Services Act (DSA) by the European Parliament (European Commission, 2022), which requires more transparency and user-adjustability in apps. Digital nudges, as reviewed by Bergram et al. (2022), while aiming to guide people's choices, they also allow users to retain autonomy by providing options and provide an avenue towards in-app self-regulation. Boosts, on the other hand, foster people's competence to make their own choices—that is, to exercise their own agency (Hertwig & Grüne-Yanoff, 2017).[3]

Business goals achieved through behavioral design in apps seem to contradict responsible design principles. Thus, our research question examines whether app design can reconcile business objectives with corporate responsibility. In our experiment, we explore an app design that includes typical behavioral design elements—aimed at fulfilling primary business goals—while also

[3] See also Fogg (2002) who pioneered a discussion on requirements for responsible behavioral design. For more details, see Table B1 in Appendix B.

providing users with choices—clearly explained and easily adjustable features—to satisfy corporate responsibility aims. We compare this new *Choice* design to a *Baseline* version with minimal functional design and a version with full *Behavioral Design*.

A related study by Mariotti et al. (2022) explores the optimal design of information nudges to influence present-biased decision-making under uncertainty. Their model predicts that targeted nudges, such as traffic-light labels, guide consumers by making the consequences of their choices more immediate and tangible. The authors emphasize the potential benefits of reduced transparency in managing impulsive behavior. In the digital context of our study, we evaluate the need for convenient tools to help users self-regulate their potentially misguided behavior and explore how much value they would place on these tools.

Our study aims to show the implications of in-app user choice, filling critical gaps in the literature. First, while previous research predominantly focuses on the effects of traditional behavioral nudges, our investigation specifically examines the holistic effects of behavioral design systematically aligned to meet the goal of increasing engagement (usage time). Second, we evaluate the potential financial implications for app providers by measuring hypothetical willingness to pay (WTP), an aspect that has received limited attention in existing studies on digital nudging. WTP reflects users' valuation based on their app experience, including aspects of empowerment and convenience that depend on the design. Additionally, our study empirically tests the effects of user choice in a practical setting—a framed field experiment with a news app designed by the authors. This innovative method not only contributes to the understanding of how customizability influences user behavior but also aids in identifying potential pathways for responsible app design and enhances the discourse around responsible app provision.

The experimental details are as follows: 141 student participants had access to the experimental news app for two weeks. There were three experimental conditions: *Baseline*, an app version featuring a minimal functional design; *Behavioral Design*, incorporating behavioral design elements; and *Choice*, which included the same behavioral design elements with additional explanatory information and adjustable design features. We observed participants' usage behavior and collected additional information through a post-study survey. We found that *Behavioral Design* resulted in an increase in usage time by 0.51 of standard deviation compared to the *Baseline* and that the usage time in *Choice* was in-between the two other treatments. We also found that the WTP for the app was lowest in the *Behavioral Design* treatment. Thus, we delivered first evidence that behavioral design combined with choice has a potential to improve business practices and enhance users' agency. However, due to the limitations of our study, those findings need further validation.

While our study's two-week timeframe limits our ability to capture digital addiction, the lower WTP observed in the *Behavioral Design* treatment may suggest that users are becoming aware of the addictive aspects associated with this design. Additionally, we acknowledge that the welfare effects of increased usage time in the context of reading news, as studied here, are not entirely clear.[4] Other limitations include the modest sample size and the use of student participants. Nonetheless, we demonstrate how dramatically consumption patterns of the same news content can change when presented in different app versions, as well as how various choice features can further modify these patterns.

[4] For example, while increased news consumption can enhance information access, concerns remain regarding exposure to misinformation, the reinforcement of echo chambers, reduced attention spans, and potential increases in digital addiction.

2. EXPERIMENTAL DESIGN

The study app was built by a professional app developer based on our concept. Specifically, the app was a news app, because they are commonly used, news feeds are an easy way to fill a study app with rich real-world and up-to-date content, and readers worldwide predominantly obtain their news content online (Flaxman et al., 2016). News came from 15 different German language sources, comprising the major news providers. However, our study does not specifically focus on news content. Importantly, while our study app is 'neutral'—not monetizing usage time (no ads) and not aiming to influence individual word perceptions (no fake or biased news)—its tempting behavioral design and the choice version could equally well be adapted to other environments, making our results generalizable. For more details on the experimental design and its mechanisms and risks see Table A1 and for essential pre-studies Tables B2 and B3 in the Appendix.

Three app versions were designed, one for each treatment group: (1) *Baseline* with a minimal functional design, but no additional behavioral design elements aiming to raise user engagement and no choice features. (2) *Behavioral Design* with behavioral design elements intending to prolong usage time but no choice features. (3) *Choice* with the same elements as those in the *Behavioral Design* version but including upfront information on pros and cons of behavioral design and offering adjustable design features. Those features included several separate options to monitor, adjust, or deactivate news filtering, and to adjust or deactivate push notifications. See Table 1, and for more details of the app versions and screenshots Table A3 and C2 in the Appendix.

We recruited 141 participants from the student pool of the Laboratory for Economic Research at Otto von Guericke University in Magdeburg for a 14-day online study conducted in November 2021. Table 2 presents participant demographics. The study was conducted

anonymously. Participants were informed that the aim of the study was to learn about their user experience with the app. All participants signed an informed consent declaration.[5] We randomly assigned the participants to three treatment groups, stratifying on gender.[6] The post-study survey was filled in completely by 136 participants, with 45-46 participants in each group, which forms our analysis sample.

Participants who completed all parts of the study received 50 euros. The app usage time was not incentivized, ensuring the natural observation of individual choices regarding time spent with the app. Beyond the direct financial incentives, the app offered free access to daily news content that is either not available without subscription or only in limited quantities. We communicated with participants via the lab and on every fourth day sent reminder emails to those who had not opened the app for the previous three days.

Our primary outcome variable is usage time, measuring user engagement and reflecting the common monetization goal of app providers. We also measured non-incentivized willingness to pay (WTP) as an indicator of users' experience of app usage. In the post-study survey, we asked each participant to state the amount of money they would be willing to pay for a comparable market-ready app in four most common business models that rely on direct payments: individual (i.e. not shared) monthly subscription, monthly subscription shared with family and friends, donation (pay-as-you-want), and paying per article (pay-as-you-use).

[5] While our university had no ethics committee at the time, we were pointed to the precedential decision on ethical approvals by the German Association for Experimental Economic Research (GfeW) which states that by German federal law and by the ethics code of the German National Science Foundation (Deutsche Forschungsgemeinschaft DFG), human subject experiments are exempt from the IRB review as long as standard experimental protocols are used. Throughout the app development and experiment design, we adhered to ethical standards and data protection requirements, also obtaining the necessary permissions from the news providers.

[6] Table 2 confirms that our randomization achieved a good balance between treatment groups. None of the 18 p-values from the test of proportions is significant at conventional levels thus confirming a good balance. However, given a small sample, some important differences might go undetected. Therefore, in the robustness checks, we include controls for those individual characteristics.

We expect the app usage time to be higher in the Behavioral Design group compared to the Baseline and Choice groups. Without choice, behavioral design exploits behavioral biases by serving users' immediate pleasures through gamification and social comparison with other readers, or by keeping them trapped in echo chambers. The usage time in the Choice group is expected to lie between that of the other two groups as it allows for self-determined active choices. This leads us to:

Hypothesis 1: *Usage time: Baseline < Choice < Behavioral Design*

We expect WTP for the app to be higher in the Choice group, compared to the Behavioral Design group. The educational boosts and adjustable design features offer a larger choice set that is more likely to meet user preferences, and thus should increase users' utility on average. In addition, there is also value added to the presentation of the choice set per se, which should lead users to make better informed choices, increasing user experience, and thus their valuation. Moreover, transparency has already been shown to increase WTP in an online privacy context (e.g. Tsai et al. 2011). This leads us to:

Hypothesis 2: *Willingness to pay: Behavioral Design < Choice*[7]

3. RESULTS

Over the 14 days of the experiment, 7,839 articles were available in the news feed, of which 2,301 (29%) were read by at least one participant. We classified an article as read if the participant opened the article for at least five seconds. The average app usage time was 6.6 minutes per day. This compares to eleven minutes spent daily by a representative German person on reading print newspapers and to three minutes spent daily on reading newspaper content online in 2021, the year of the study. However, the usage time differs by group, see Table 3. Participants in the *Baseline*

[7] We have no prior on the WTP for *Baseline* in comparison to other groups.

group spent an average of 67 minutes over the two weeks. This corresponds to slightly less than five minutes per day. In the *Behavioral Design* group, the average usage time was 132 minutes, which is almost ten minutes per day. In the *Choice* group, it was 88 minutes, which is over six minutes per day. The time spent on reading articles alone was 39, 86, and 54 minutes in the respective treatments, which is between close to three minutes per day in the *Baseline* group to more than six minutes per day in the *Behavioral Design* group. Thus, the observed usage time in the *Behavioral Design* group is double that of the *Baseline* group.

To reduce the number of hypothesis tests, we combined the measures of total time usage and time spent reading articles using principal component analysis. The resulting standardized variable facilitates the interpretation of effect sizes in the regressions. Table 4, Column I, presents results of an OLS regression controlling for the stratification variable (gender).[8] It suggests a 0.17 of standard deviation higher usage time in the *Choice* treatment than in the *Baseline*, significant at $p<0.1$ (two-sided test). *Behavioral design* increases usage time by as much as 0.51 of standard deviation ($p<0.05$). Both differences turn insignificant once we account for multiple hypothesis testing (MHT) using the Romano-Wolf method (Clarke 2021) but, on the other hand, randomization p-value for Westfall-Young multiple testing of treatment significance (Young 2019), rejects the hypothesis of no overall treatment effects. When comparing *Choice* and *Behavioral Design*, a Wald test in Column I yields insignificant differences between those treatments. Applying a nonparametric Wilcoxon rank-sum (Mann-Whitney) test to pairwise comparisons rejects equality between *Choice* and *Baseline* ($p=0.079$), *Baseline* and *Behavioral Design* ($p=0.032$) but not between *Behavioral Design* and *Choice* ($p=0.686$). Overall, while the

[8] Table A4 in the Appendix shows that the coefficients remain stable when including further control variables.

time use averages are ordered in line with our hypothesis 1, the differences between the *Behavioral Design* and *Choice* groups are not statistically different.

Since the averages may mask the dynamics of the usage time, we also shortly report usage patterns over time. In Figure 1, we present selected outcome variables by treatment group and over time. The graphs for additional variables can be found in Figure A1 in Appendix A. The first graph in Figure 1 shows the development of the article reading time over the study period for the three groups separately. We see that, at the beginning of the study, the reading time per day in the *Choice* group was more similar to the *Behavioral Design* group than to the *Baseline* group and from day seven on it was the other way around. We observe a more negative trend in usage over time in the *Choice* group than in the two other groups.[9]

To better understand user engagement in the *Choice* group, we briefly report on the actual usage of choice features in the *Choice* group—the only treatment group with such features (see also Figure A3 in the Appendix). We first note that the majority of the participants (66%) changed the offered filtering settings (e.g. deactivation, displaying popular articles first, etc.) including the adjustment of the filtering algorithm's outcome: 30% did so once and 36% did so multiple times. Overall, the participants adjusted filtering options throughout the study, including on the last day. Altogether, 74% of the participants adjusted at least one of the default settings. The high usage rate of choice features, when offered so, confirms their value for users, and we might observe a learning or habituation effect over time.

Next, we asked the participants for their WTP for an individual monthly subscription, shared subscription, the amount they would be willing to donate, or how much they would be

[9] While our treatment does not allow us to distinguish between the effects of transparency and choice, the observed pattern suggests a stronger impact of choice—individuals appeared to learn over time—whereas transparency, presented upfront, seems to have had no effect.

willing to pay per article. Table 3 below shows the averages of the respective outcome variables. The WTP is lowest in the *Behavioral Design* across all measures. Once again, to streamline hypothesis testing, we combined all measures into a single variable using principal component analysis. The resulting standardized variable simplifies the interpretation of effect sizes in the regressions presented in Table 4, Column II. The *Choice* and *Baseline* groups have similar WTP, while the WTP in the *Behavioral Design* is significantly lower. While accounting for the MHT does not reject the similarity of treatments, the Westfall-Young multiple testing procedure does reject the hypothesis of no overall treatment effect. Comparing the *Choice* and *Behavioral Design* groups, a Wald test reveals a significant difference between the treatments ($p < 0.05$). Additionally, a nonparametric Wilcoxon rank-sum (Mann-Whitney) test for pairwise comparisons rejects the equality between the *Behavioral Design* and *Choice* groups ($p = 0.047$), confirming our Hypothesis 2. The same is true for the comparison between the *Baseline* and *Behavioral Design* groups ($p = 0.045$).

4. DISCUSSION AND CONCLUSIONS

Odysseus' story about avoiding the Sirens' behavioral control over his free choice illustrates the problem of tempting apps and the solution in form of prior awareness raising and, based on that, informed decision-making and self-regulation, as it is offered in the choice version of the study app. Nevertheless, the choice version does not kill pleasure, since it is convenient to use and still offers the entire range of behavioral design elements. Just like Odysseus still enjoyed listening to the Sirens while having taken care to protect himself.

Our two-week-long field experiment with a news app provides causal evidence that behavioral design significantly increases app use and lowers WTP compared to the baseline with a minimal functional design. Offering consumers choice in the version with explanatory

information and adjustable design features leads to ‘intermediate’ usage time and significantly higher WTP than in the version with behavioral design. Specifically, we measured engagement through usage time and verified that behavioral design doubles average usage time compared to the baseline, aligning with Luguri’s & Strahilevitz’s (2021) findings on behavioural design effects in real-world apps. Choice empowers users to make better decisions about how much to engage with a tempting app. On average, users value having alternative choices in apps more than having a full behavioral design. Thus, this study contributes to the understanding of the time-increasing effect of concerted behavioral design in apps and the valuation of transparency and choice measures (Cemiloglu et al., 2023; European Commission, 2022; Mariotti et al. 2022). However, these findings should be interpreted with caution due to the limitations of this study, including a small student sample size and a limited scope of measurements. Further research is necessary to replicate and expand upon these results.

The welfare implications of reduced app usage in the context of a news app remain ambiguous. While increased reading time can benefit users when the news is personally relevant and meaningful (Schrøder, 2019), it poses risks if users are merely driven to compete with others, read too quickly without adequate attention, or become trapped in echo chambers. Moreover, real apps may pose additional risks, such as digital addition or spread of fake news. There is also the potential misuse of behavioral data collected during usage, all which we could not replicate in our study for ethical reasons. In this regard, the trustworthy academic integrity of our study represents a limitation, and the measured WTP may be inflated.

Regarding external validity of our results, we believe that they are generalizable to any app, because behavioral design that intends to increase usage time and succeeds even for a rather serious and common everyday activity like news reading, is also likely to succeed in other serious or more

playful, mundane or extraordinary environments too. Whether behavioral design elements are "beneficial" or "dark", depends on the targeted user behavior and the intention of the designer (Thaler, 2018). Whether the design is tempting, will depend on what individual users fall for. Our study merely showcases what systematically concerted behavioural design and a range of choice features can achieve, and what effect size they can have, independently from any business model or further context. Thus, responsible design does not necessarily implement making any normative decisions on behalf of the users.

By the design and use of our authentic study app—a pioneering research method—we opened up the way for future complex app studies in controlled environments. For example, future studies might investigate more details of the underlying mechanisms of single behavioral design elements or social network effects. For this to succeed, a large user pool is required, either by an own testbed of a commercially run study app or jointly with providers of real apps.

Our findings have important implications for app providers and policymakers. From a corporate social responsibility (CSR) perspective, offering users transparent and adjustable design features could represent a shift toward more ethical app development practices. Such practices not only enhance user satisfaction but may also foster trust and long-term customer loyalty. For practitioners in the field of digital media, our practical contribution is also to show an example and provide a first documented, comprehensible, feasible, and reproducible proof of concept for the design of a responsible app with regard to its behavioral design and its risks. The monitoring and correction of algorithmic news filtering results even provides an example for transparent interaction with algorithms, which is deemed a valuable insight for software engineers, as we derive from their feedback in interviews. The reproducibility of our design may be applied in any

app, scaling the extent of our practical impact. Therefore, our research insights have the potential to lead to groundbreaking future trends.

The managerial implications of our research are important too. The changing regulatory conditions for app design open up new possibilities for new business models and market positioning, as well as new industry standards, for which we provided a practical tool, verified by methods of behavioral economics.

Table 1: Three app versions

BASELINE 

No behavioral design that raises user engagement

- In-app consent to privacy policy and terms & conditions, optional media competency quiz
- **Minimal functional design:**
  - Article rating and display of average article ratings
  - Display of number of readers
  - Same grey design for all news categories in the news feed
- No additional behavioral design elements to raise user engagement
- No choice features

BEHAVIORAL DESIGN 

Full behavioral design that raises user engagement

- In-app consent to privacy policy and terms & conditions, optional media competency quiz
- **Full behavioral design** that would reasonably be added in a real app:
  - Personalized algorithmic news filtering based on news categories
  - Choice architecture: distinct colors marking news categories and highlighted news
  - Gamified elements: achievement badges, rankings, progress history
  - Tailored push notifications on news updates, rewards, and ranking
- No choice features

CHOICE 

Adjustable design features & information on chances & risks

- In-app consent to privacy policy and terms & conditions, optional media competency quiz
- **In-app consent to behavioral design** (with information on chances & risks)
- Same behavioral design elements as the Behavioral Design version
- **Adjustable design features** with short local explanations:
  - Monitor, adjust, or deactivate news filtering (several options available)
  - Adjust or deactivate push notifications

Table 2: Balancing table

	Baseline (A)	*Behavioral Design (B)*	*Choice (C)*	Test of proportions p-value		
		Share		A=B	A=C	B=C
Female	0.533	0.556	0.522	0.832	0.912	0.746
Native speaker	0.956	1.000	0.978	0.153	0.544	0.320
With no reading problems (impaired vision, reading difficulties)	0.622	0.578	0.457	0.667	0.113	0.247
Aged 18-25	0.556	0.511	0.500	0.673	0.596	0.916
With a BA degree	0.356	0.378	0.500	0.827	0.164	0.240
With a MA degree	0.244	0.244	0.130	1.000	0.163	0.163
Computer science major	0.133	0.111	0.109	0.748	0.718	0.971
N	45	45	46			

Table 3: Means of the usage time and willingness to pay (WTP) variables

			Usage time		WTP			
	N		Total usage time	Reading time	Monthly subscription	Monthly shared subscription	Donation	Per article
Baseline	45	Mean	66.784	39.709	5.167	9.856	5.633	0.867
		Std. error	(8.855)	(5.546)	(0.616)	(1.110)	(0.982)	(0.442)
Behavioral Design	45	Mean	132.009	85.677	3.811	7.233	4.478	0.328
		Std. error	(30.308)	(20.638)	(0.480)	(0.898)	(0.853)	(0.075)
Choice	46	Mean	87.526	54.472	4.793	9.554	6.054	0.835
		Std. error	(9.553)	(6.637)	(0.462)	(1.029)	(0.981)	(0.339)
Baseline= Behavioral		Two-sided t-test p-value	0.044	0.036	0.086	0.070	0.377	0.236
Baseline= Choice			0.167	0.156	0.144	0.093	0.228	0.150
Behavioral =Choice			0.115	0.091	0.629	0.843	0.762	0.955

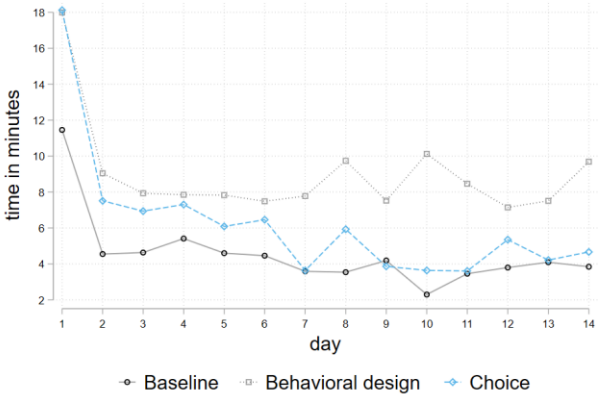
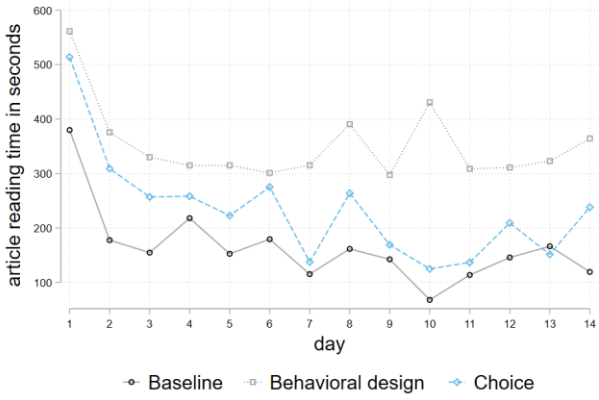
Notes: Total usage time is measured as time between opening and closing the app. Reading time is the time spent on reading articles—the article is classified as read if it is opened for at least five seconds. We also cap the maximum reading time at approximately 3-4 minutes, which is the time that an average reader would need to read the text. The WTP variables were measured in the survey at the end of the experiment. The questions read (translation): (1) *How much should a monthly subscription for a comparable market-ready newsfeed reader app cost?* (2) *How much should a monthly subscription for a comparable market-ready newsfeed reader app cost if it can be shared with family and friends?* (3) *How much should a monthly donation be for the provider of a comparable, free of cost, market-ready newsfeed reader app?* The choice was made on a slider with values up to 50 euros. The last question was: (4) *How much should one article read in a comparable, market-ready newsfeed reader app cost?* followed by a write-in answer in euros.

Table 4: The effects of treatments on usage time and willingness to pay (WTP)

	Usage	WTP	Column I+II
	I	II	III
Choice	0.172 (0.081) [0.227]	-0.029 (0.898) [0.902]	
Behavioral Design	0.510 (0.036) [0.227]	-0.418 (0.055) [0.227]	
Randomization p-value for Westfall-Young multiple testing of treatment significance	0.024[a]	0.076[a]	0.047[b]
Wald test p-value for *Choice=Behavioral Design*	0.162	0.022	
Observations	136	136	
R2	0.058	0.089	

Notes: OLS regression with strata fixed effects (gender); Outcome variables: To reduce the number of hypothesis tests, we combined the measures of time (total usage time and time spent reading articles) and the measures of WTP (monthly subscription, monthly shared subscription, donation, per article) into one dependent variable each using principal component analysis. The resulting variables are standardized, which facilitates the interpretation as effect sizes; p-values in parentheses are derived from robust standard errors; square brackets contain Romano-Wolf adjusted p-values for multiple testing (two regressions and two treatments in each). Westfall-Young p-value from: [a] the test that there is no effect of any treatment on the outcome, and [b] the test of the null of complete irrelevance (no treatment had no effect on any outcome).

Figure 1: Dynamics of average usage time per study day and by treatment groups

A: Total usage time in minutes	B: Article reading time in seconds
	

5. REFERENCES

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. Science, 347(6221), 509–514. doi.org/10.1126/SCIENCE.AAA1465

Acquisti, A., Taylor, C., & Wagman, L. (2016). The Economics of Privacy. Journal of Economic Literature, 54(2), 442–492. doi.org/10.1257/jel.54.2.442

Allcott, H., Braghieri, L., Eichmeyer, S., & Gentzkow, M. (2020). The welfare effects of social media. American Economic Review, 110(3), 629–676. doi.org/10.1257/aer.20190658

Allcott, H., Gentzkow, M., & Song, L. (2022). Digital addiction. American Economic Review, 112(7), 2424-2463. 10.1257/aer.20210867

Alter, A. (2017). Irresistible: The rise of addictive technology and the business of keeping us hooked. Penguin.

Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Fallin Hunzaker, M. B., Lee, J., Mann, M., Merhout, F., & Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. Proceedings of the National Academy of Sciences of the United States of America, 115(37), 9216–9221. doi.org/10.1073/pnas.1804840115

Beam, M. A. (2014). Automating the news: How personalized news recommender system design choices impact news reception. Communication Research, 41(8), 1019–1041. doi.org/10.1177/0093650213497979

Bergram, K., Djokovic, M., Bezençon, V., & Holzer, A. (2022). The digital landscape of nudging: A systematic literature review of empirical research on digital nudges. CHI Conference on Human Factors in Computing Systems (CHI '22). doi.org/10.1145/3491102.3517638

Boerman, S. C., Kruikemeier, S., & Zuiderveen Borgesius, F. J. (2017). Online Behavioral Advertising: A Literature Review and Research Agenda. Journal of Advertising, 46(3), 363–376. doi.org/10.1080/00913367.2017.1339368

Cemiloglu, D., Arden-Close, E., Hodge, S. E., & Ali, R. (2023). Explainable persuasion for interactive design: The case of online gambling. Journal of Systems and Software, 195, 111517. doi.org/10.1016/j.jss.2022.111517

Clarke, D., 2021. rwolf2 Implementation and Flexible Syntax. http://www.damianclarke.net/computation/rwolf2.pdf

Claussen, J., Kretschmer, T., & Mayrhofer, P. (2013). The Effects of Rewarding User Engagement: The Case of Facebook Apps. Information Systems Research, 24(1), 186–200. doi.org/10.1287/isre.1120.0467

Combs, T. D., & Brown, R. A. (2018). Digital behavioral design. In Venice Beach: Boundless Mind.

Dellarocas, C. (2006) Strategic Manipulation of Internet Opinion Forums: Implications for Consumers and Firms. Management Science 52(10):1577-1593. doi.org/10.1287/mnsc.1060.0567

European Commission, Directorate-General Consumers, Lupiáñez-Villanueva, F., Boluda, A., Bogliacino, F., Liva, G., Lechardoy, L., & de las Heras Ballell, T. (2022). Behavioural study on unfair commercial practices in the digital environment : dark patterns and manipulative personalisation : final report. Publications Office of the European Union. doi.org/doi/10.2838/859030

Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. Public Opinion Quarterly, 80(S1), 298–320. doi.org/10.1093/poq/nfw006

Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2018). Human Decisions and Machine Predictions. The Quarterly Journal of Economics, 133(1), 237–293. doi.org/10.1093/QJE/QJX032

Lambrecht, A., Goldfarb, A., Bonatti, A., Ghose, A., Goldstein, D. G., Lewis, R., Rao, A., Sahni, N., & Yao, S. (2014). How do firms make money selling digital goods online? Marketing Letters, 25(3), 331–341. doi.org/10.1007/s11002-014-9310-5

Luguri, J., & Strahilevitz, L. J. (2021). Shining a Light on Dark Patterns. Journal of Legal Analysis, 13(1), 43–109. doi.org/10.1093/jla/laaa006

Mariotti, T., Schweizer, N., Szech, N., & von Wangenheim, J. (2022). Information Nudges and Self-Control. Management Science, 69(4), 2182-2197. doi.org/10.1287/mnsc.2022.4428

Mathur, A., Acar, G., Friedman, M. J., Lucherini, E., Mayer, J., Chetty, M., & Narayanan, A. (2019). Dark Patterns at Scale. Proceedings of the ACM on Human-Computer Interaction, 3(CSCW), 1–32. doi.org/10.1145/3359183

Michie S., Richardson M., Johnston M., Abraham C., Francis J., Hardeman W., Eccles M. P., Cane J., and Wood, C. E. (2013), The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions, Annals of behavioral medicine, 46 (1), 81-95. doi: 10.1007/s12160-013-9486-6

Montag, C., Lachmann, B., Herrlich, M., & Zweig, K. (2019). Addictive Features of Social Media/Messenger Platforms and Freemium Games against the Background of

Psychological and Economic Theories. International Journal of Environmental Research and Public Health, 16(14), 2612. doi.org/10.3390/IJERPH16142612

Mosquera, R., Odunowo, M., McNamara, T., Guo, X., & Petrie, R. (2020). The economic effects of Facebook. Experimental Economics, 23(2), 575–602. doi.org/10.1007/s10683-019-09625-y

Oinas-Kukkonen, H., & Harjumaa, M. (2008). A systematic framework for designing and evaluating persuasive systems. In: Oinas-Kukkonen, H., Hasle, P., Harjumaa, M., Segerståhl, K., Øhrstrøm, P. (eds) Persuasive Technology. PERSUASIVE 2008. Lecture Notes in Computer Science, vol 5033. doi.org/10.1007/978-3-540-68504-3_15

Schrøder, K. (2019). What do news readers really want to read about? How relevance works for news audiences. Digital News Project. Reuters Institute for the Study of Journalism. doi.org/10.60625/risj-n12y-az27

Susser, Daniel, Roessler, Beate and Nissenbaum, Helen F., (2019). Online Manipulation: Hidden Influences in a Digital World. 4 Georgetown Law Technology Review 1 dx.doi.org/10.2139/ssrn.3306006

Thaler, R. H. (2018). Nudge, not sludge. Science 361, 431-431. doi.org/10.1126/science.aau9241

Tsai, J. Y., Egelman, S., Cranor, L., & Acquisti, A. (2011). The effect of online privacy information on purchasing behavior: An experimental study. Information systems research, 22(2), 254-268. doi.org/10.1287/isre.1090.0260

Young, Alwyn (2019). Channelling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results. Quarterly Journal of Economics 134 (2): 557-598. doi.org/10.1093/qje/qjy029

Online Appendix for:
The Impact of Behavioral Design and Users' Choice on Smartphone App Usage and Willingness to Pay: A Framed Field Experiment

APPENDIX A: DETAILS OF THE EXPERIMENTAL DESIGN AND ADDITIONAL ANALYSIS

Table A1: Details of the experimental design, describing mechanisms and risks

Experimental design
App versions
The app versions in the *Behavioral Design* and *Choice* groups contained behavioral design elements typically found in similar real apps. These elements included algorithmic filtering of news based on user preference profiling using the news categories, choice architecture (e.g. different coloring of news categories and highlighted news), gamified elements (e.g. rankings and achievement badges for the most active readers), tailored push notifications on rewards and latest news, and user-friendly and attention-grabbing front-end elements to make the app more engaging. Each of these design elements poses risks for users. Algorithmic filtering and choice architecture may inadvertently trap users in echo chambers. Gamified elements, by leveraging neurotransmitter rewards, hold the potential for causing addictive reactions. Push notifications not only trigger the collection of rewards but also spur news reading, and contribute to a fear of missing out on the latest news. Finally, the attention-grabbing nature and user-friendliness of interfaces may inadvertently lead users towards an increasingly passive and dulled consumption of news. The *Choice* version additionally contained novel consumer protection measures addressing the above-named risks. The protection measures consisted of (1) explanations of the design mechanisms and their risks in form of an in-app consent to behavioral design and (2) features that enabled participants to adjust design settings. The upfront information (i.e. in-app consent, not to confuse with the informed consent to the experiment) read as follows: "*New habits, automation of recurring processes, automated decisions and a digital environment tailored to your needs simplify your life, but also invisibly influence your behavior. Here, such behavioral design is transparent and adjustable. Keep control at all times! You can change the settings at any time. We have pre-selected everything for full experience, optimal functionality, and maximum fun.*" The adjustable settings allowed participants to tailor the functionalities of their app and to deactivate certain behavioral design elements.

News and design around news

News came from 15 different German language sources, comprising the major news providers, from whom we gathered the necessary permissions and licenses. Our app did not contain any advertising; thus, there was no actual monetization of users' time. To unify the design, we only included one image per article.

News items were assigned to one of four predefined categories in the feed. The categories were (1) economy and politics, (2) sport, (3) culture and entertainment, and (4) gaming and technology. The category names and the highlighting were grey in the *Baseline Design* version and had distinct colors in the other two groups.

The content in each participant's news feed was automatically updated several times each day. Additionally, participants in the *Behavioral Design* and *Choice* groups could update their news feed through the widely used "pull-to-refresh" design element, which allows users to load new content by "pulling down" the screen with one finger.

In the main news feed, all app versions displayed the average article rating and the number of people who read an article. The option to rate an article (i.e. how worthy it is to read) was also available in all app versions at the end of each article on a voluntary basis. This minimal behavioral design in the *Baseline* group aimed to make the app somewhat realistic and allowed for meaningful responses in the post-study survey.

Participant payoffs

Participants who completed all parts of the study received 50 euros. The fee was divided into 6 euros for downloading the app, 1 euro for each day participating in the study, without any requirement to actively use the app, and 30 euros for filling in the post-study survey, aiming to reduce attrition. The app usage time was not incentivized.

Reminder emails and control questions

The reminders read *"You have not used the news app since several days. During the study duration you may freely use the functions of the app and read and rate the content as you like."* With the reminder emails we tried to avoid complete inactivity, while having the least impact on app usage time. On top of that, using control questions in the informed consent declaration, participants were aware that they could stay inactive during the whole study, even if they receive reminders. In total, including those who did not complete the post-study survey and were excluded from the analysis, we sent 57 reminders to 39 participants. The numbers were 28 and 18 in the *Baseline*, 15 and 10 in the *Behavioral Design*, and 14 and 11 in the *Choice*.

Withdrawals

Although withdrawal was always possible with a proportional payment until the day of withdrawal, there were no active withdrawals.

Table A2: Media competency quiz

Media competency quiz
Description
Each of the three app versions contained an optional media competency quiz. This quiz was placed saliently into the user account and in-app consent settings, which were the first settings participants had to make before starting to use the app. The in-app consent settings included privacy policy, terms of usage, and for the *Choice* group agreement to behavioral design. Later, during the whole study, the quiz was available under the settings menu. Filling in the quiz was not obligatory. However, participants had to choose whether they gave their in-app consent with or without completing the quiz.
Aim of the quiz
We explained to participants that the quiz could be done to measure one's media competency and upon that to decide whether one can make the decision to agree to the in-app consent (1) quickly, flying over the text, or (2) carefully, reading the text line-by-line. In the *Baseline* and *Behavioral Design* groups the quiz contained only questions related to privacy and data protection. In the *Choice* group, it additionally contained questions related to behavioral design. The reason for building in the quiz was to see if there's a need to fetch different kinds of users offering them distinguished ways to make their in-app consent. With this, we aimed to make the in-app consent more inclusive. Users might differ in their knowledge and experiences with privacy and data protection and protection from the influence of behavioral design, or they might have time constraints, or varying motivations. When presenting settings and explanations, we tried to consider how much time and what kind of context participants need for making their protection decisions. Participants thus had the option to quickly read through the explanations and orient with the help of icons and a clear structure of the presented information, or they could go into details of explanations in simple language by opening accordion folds.
Results
45% of the participants agreed to the in-app consent quickly, flying over the text and 55% carefully, reading the text line-by-line. Although the media competency quiz was started and finished by only 28% of the participants, 81% of all participants who filled in the post-study survey found it to be useful.

Table A3: Exemplary screenshots of the three app versions

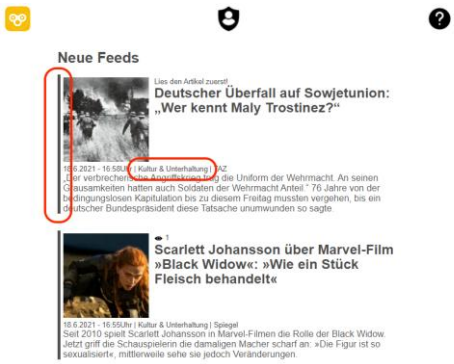
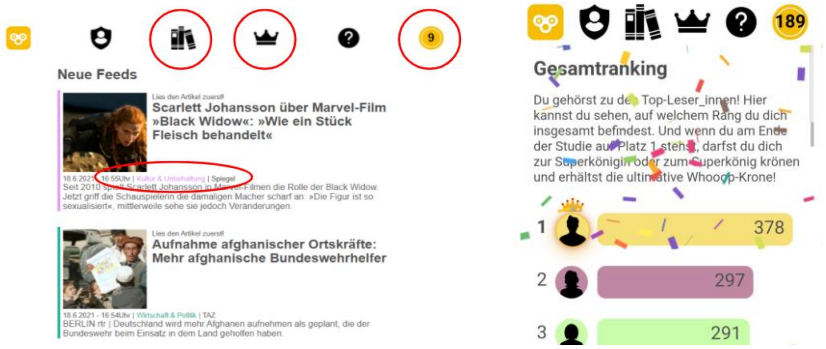
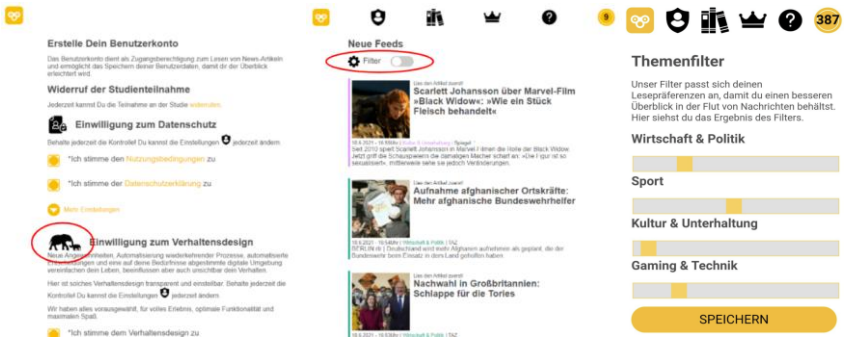
Baseline	
Grey marking of all news categories in the news feed	
Behavioral Design	
Choice architecture: distinct colors marking news categories Gamified elements accessible through the menu: achievement badges, rankings, history for tracking progress	
Choice	
In-app consent to behavioral design News filtering can be monitored, adjusted, or deactivated	

Table A4: The effects of treatments on usage time: robustness checks

	Usage	Usage	total time usage	total time usage	time spent reading articles	time spent reading articles
Choice	0.172	0.172	21.711	21.457	15.362	15.580
	(0.081)	(0.136)	(0.096)	(0.156)	(0.076)	(0.124)
Behavioral design	0.510	0.479	64.579	60.554	45.568	42.746
	(0.036)	(0.032)	(0.040)	(0.037)	(0.033)	(0.029)
Native German speaker		0.327		42.487		28.445
		(0.226)		(0.185)		(0.278)
Reading restrictions		0.090		9.653		9.171
		(0.586)		(0.657)		(0.519)
Age group: 18 to 25		-0.504		-61.792		-46.272
		(0.042)		(0.059)		(0.029)
Bachelor degree		-0.275		-32.645		-26.013
		(0.162)		(0.208)		(0.124)
Master degree		-0.175		-20.660		-16.564
		(0.540)		(0.580)		(0.503)
Share computers science major		-0.110		-10.900		-11.817
		(0.523)		(0.651)		(0.402)
Constant	-0.227	-0.161	66.670	72.452	39.639	47.292
	(0.000)	(0.650)	(0.000)	(0.106)	(0.000)	(0.148)
Observations	136	136	136	136	136	136
R^2	0.058	0.131	0.057	0.121	0.058	0.140

Notes: OLS regression with strata fixed effects (gender); p-values in parentheses are derived from robust standard errors; Usage combines the measures of total time usage and time spent reading articles using principal component analysis.

Table A5: The effects of treatments on WTP: robustness checks

	WTP	WTP	WTP for monthly subscription	WTP for monthly subscription	WTP for monthly shared subscription	WTP for monthly shared subscription	Willing to donate	Willing to donate	WTP per article read: yes	WTP per article read
Choice	-0.029	-0.053	-0.319	-0.368	-0.161	-0.607	0.467	0.748	-0.028	-0.081
	(0.898)	(0.807)	(0.677)	(0.627)	(0.912)	(0.679)	(0.737)	(0.598)	(0.961)	(0.863)
Behavioral design	-0.419	-0.411	-1.391	-1.404	-2.716	-2.915	-1.187	-0.749	-0.542	-0.530
	(0.055)	(0.055)	(0.073)	(0.070)	(0.056)	(0.048)	(0.360)	(0.550)	(0.226)	(0.180)
Native German speaker		-0.131		0.497		2.186		-7.641		0.111
		(0.590)		(0.612)		(0.349)		(0.177)		(0.828)
Readin		-0.031		-0.025		-0.672		0.285		-0.039

g restrictions		(0.859)		(0.968)		(0.574)		(0.813)		(0.922)
Age group: 18 to 25		0.132 (0.478)		0.306 (0.645)		-0.506 (0.680)		2.271 (0.079)		0.224 (0.619)
Bachelor degree		0.244 (0.230)		0.628 (0.388)		1.038 (0.441)		0.537 (0.678)		0.714 (0.125)
Master degree		0.032 (0.846)		0.264 (0.682)		-0.741 (0.607)		0.437 (0.764)		0.168 (0.417)
Share computers science major		0.117 (0.793)		0.470 (0.734)		-1.074 (0.574)		-0.774 (0.628)		1.104 (0.357)
Constant	0.148 (0.422)	0.109 (0.681)	5.160 (0.000)	4.180 (0.000)	9.839 (0.000)	8.404 (0.001)	5.628 (0.000)	11.298 (0.052)	0.866 (0.055)	0.218 (0.648)
Observations	136	136	136	136	136	136	136	136	136	136
R^2	0.089	0.108	0.078	0.088	0.122	0.138	0.023	0.081	0.014	0.067

Notes: OLS regression with strata fixed effects (gender); p-values in parentheses are derived from robust standard errors; WTP combines the four measures of WTP using principal component analysis.

Table A6: Further app usage statistics

	N		Number of articles screened	Number of articles read	Number of articles rated	Number of sessions	Share using the app per day
Baseline		Mean	28.133	23.311	18.444	26.200	0.708
	45	Std. error	2.715	2.503	2.504	2.141	0.033
Behavioral Design		Mean	118.533	96.756	89.956	38.067	0.806
	45	Std. error	32.259	25.409	31.162	3.211	0.026
Choice		Mean	67.891	52.500	42.261	39.870	0.799
	46	Std. error	9.926	7.199	8.370	3.661	0.029

Table A7: Probability of being a top third most active user per treatment group

	All participants	Only *Behavioral Design* group
Native German speaker	0.062 (0.295)	-
Reading restrictions	0.201*** (0.075)	0.127 (0.157)
Age group: 18 to 25	-0.280*** (0.079)	-0.315** (0.155)
Bachelor degree	-0.198** (0.084)	-0.162 (0.180)
Master degree	-0.120 (0.110)	-0.134 (0.191)
Female	0.179** (0.077)	0.080 (0.148)
Computer science major	-0.214*** (0.079)	-0.357** (0.137)
Economics major	0.103 (0.086)	0.080 (0.153)
Constant	0.311 (0.321)	0.483** (0.199)
Observations	136	45
R^2	0.198	0.200

Notes: Linear probability regression (OLS); Robust errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A8: Awareness of behavioral design stated in the post-study survey

	N	share
Baseline	45	0.222
Behavioral Design	45	0.533
Choice	46	0.717

Notes: Share of participants who responded positively to the question whether they realized behavioral design aiming at increasing usage time.

Table A9: Subjective need for protection measures (among all participants)

	Share yes (N=136)
Behavioral design in general	
"*In general, what do you think of behavior-influencing elements in apps in terms…* "*Was halten Sie generell von verhaltensbeeinflussenden Elementen in Apps hinsichtlich…*	
... of usefulness?: I find it useful." *... der Nützlichkeit?: Finde ich nützlich."*	0.42
... of risk?: I find it problematic." *... der Risiken?: Finde ich problematisch."*	0.75
Behavioral algorithms	
"What do you think of behavior-influencing algorithms in apps in terms… *"Was halten Sie am ehesten von verhaltensbeeinflussenden Algorithmen in Apps hinsichtlich…*	
... of usefulness?: I find it useful." *... der Nützlichkeit?: Finde ich nützlich."*	0.50
... of risk?: I find it problematic." *... der Risiken?: Finde ich problematisch."*	0.74
Push notifications	
"What do you think most about push notifications in apps in terms... *"Was halten Sie am ehesten von Push-Nachrichten in Apps hinsichtlich...*	
... of usefulness?: I find it useful." *... der Nützlichkeit?: Finde ich nützlich."*	0.54
... of risk?: I find it problematic." *... der Risiken?: Finde ich problematisch."*	0.40
Gamification features	
"What do you think most about gamification features in apps in terms of... *"Was halten Sie am ehesten von Spielfunktionen in Apps hinsichtlich...*	
... of usefulness?: I find it useful." *... der Nützlichkeit?: Finde ich nützlich."*	0.27

... of risk?: I find it problematic." *... der Risiken?: Finde ich problematisch."*	0.54
Design elements that influence behavior *"What do you think of behavior-influencing design elements in apps in terms…* *"Was halten Sie am ehesten von verhaltensbeeinflussenden Design-Elementen in Apps hinsichtlich…*	
... of usefulness?: I find it useful." *... der Nützlichkeit?: Finde ich nützlich."*	0.62
... of risk?: I find it problematic." *... der Risiken?: Finde ich problematisch."*	0.24
No control over behavioral design: feel bad	
"How do you feel about algorithms and behavior-influencing elements making decisions about settings, push notifications, profiling, and filtering content without your control as a user?: I feel uncomfortable with this." *"Wie fühlen Sie sich damit, wenn Algorithmen und verhaltensbeeinflussende Elemente ohne ihre Kontrolle als Nutzer_in Entscheidungen über Einstellungen, Push-Nachrichten, Profilerstellung und Filtern des Inhalts treffen?: Ich fühle mich unwohl damit."*	0.69
Control over behavioral design elements: yes	
"Would you like to be able to set yourself which behavior-influencing elements are allowed to be active in the apps that you use?" *"Möchten Sie gern selbst einstellen können, welche verhaltensbeeinflussenden Elemente in von Ihnen genutzten Apps aktiv sein dürfen?"*	0.99
Regulation of behavioral design elements: yes	
"Do you think there should be legislation regulating what behavior-influencing elements can be included in apps?" *"Sollte Ihrer Ansicht nach gesetzlich geregelt werden, welche verhaltensbeeinflussenden Elemente in Apps eingebaut werden dürfen?"*	0.78
Certificates: yes	
"Do you think there should be certificates that make it transparent to users what behavior-influencing elements are built into apps?" *"Sollte es Ihrer Ansicht nach Zertifikate geben, die den Nutzer_innen transparent machen, welche verhaltensbeeinflussende Elemente in Apps eingebaut sind?"*	0.90
Easy control over data collection and use: yes	
"The collection and processing of data by companies that offer apps can be done in a variety of ways. To what extent do you agree with the following: I think that I should be given easy ways to control what data is being collected about me by companies and how it is used." *"Die Erfassung und Verarbeitung von Daten durch Unternehmen, die Apps anbieten, kann auf unterschiedliche Weise erfolgen. Inwieweit stimmen Sie den folgenden Möglichkeiten zu: Ich finde, dass ich einfache Kontrollmöglichkeiten erhalten sollte, welche Daten von Unternehmen über mich erhoben werden und wie sie verwendet werden."*	0.76
Sell data: yes	
"From my perspective, companies could be allowed to resell the data to other companies." *"Aus meiner Sicht dürften Unternehmen die Daten an andere Unternehmen weiterverkaufen."*	0.49

Share anonymized data with third parties: yes	
"I think it would be fine if companies were to pass on data collected for the purpose of market research to third parties in anonymized form." *"Ich fände es in Ordnung, wenn Unternehmen erfasste Daten zu Zwecken der Marktforschung in anonymisierter Form an Dritte weitergeben würden."*	0.23
State laws on data use: yes	
"I would like to see the state regulate and enforce by law which data may be used." *"Ich wünsche mir, dass der Staat gesetzlich regelt und durchsetzt, welche Daten genutzt werden dürfen."*	0.38
Like filters: yes	
"Some participants had the possibility to influence the order of the displayed articles: they had the choice to turn on and off several automated filters including automated filtering based on their preferences for the topics and mood of the articles.: Do you think it is good to have such a setting option in an app?" *"Einige Teilnehmer_innen hatten die Möglichkeit, die Reihenfolge der angezeigten Artikel zu beeinflussen: sie hatten die Wahl, mehrere automatische Filter an- und auszuschalten, darunter auch eine automatisierte Filterung aufgrund ihrer Präferenzen bezüglich der Themen und der Stimmung der Artikel.: Finden Sie es gut, in einer App so eine Einstellmöglichkeit zu haben?"*	0.77
Use filters: yes	
"Did you or would you have used this option in the study to influence the content that is displayed?" *"Haben oder hätten Sie diese Möglichkeit in der Studie genutzt, um den Inhalt zu beeinflussen, der angezeigt wird?"*	0.70
Like push notification settings: yes	
"Some participants were given the opportunity to set when they would receive push notifications and why they are triggered. Do you think it is good to have such a setting option in an app?" *"Einige Teilnehmer_innen hatten die Möglichkeit, selbst einzustellen, wann sie Push-Nachrichten bekommen und warum diese ausgelöst werden.: Finden Sie es gut, in einer App so eine Einstellmöglichkeit zu haben?"*	0.67
Use push notification settings: yes	
"Did you or would you have used this option in the study to influence the content that is displayed?" *"Haben oder hätten Sie diese Möglichkeit in der Studie genutzt, um den Inhalt zu beeinflussen, der angezeigt wird?"*	0.56

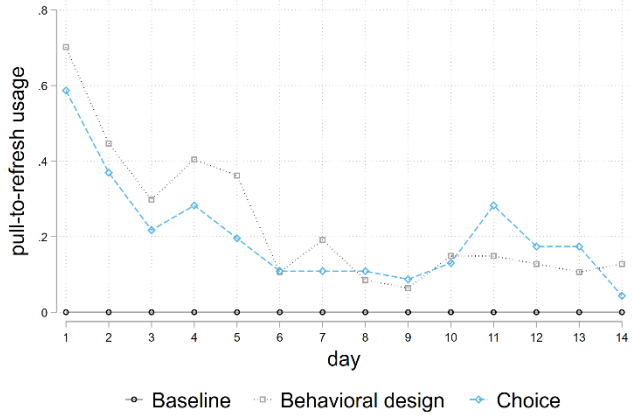
Table A10: Business models preferred by participants (multiple answers possible)

"*Which usage model would you prefer for a comparable market-ready newsfeed reader app?"* "*Welches Nutzungsmodell würden Sie für eine vergleichbare marktreife Newsfeed-Reader App bevorzugen?"*	Share yes (N=136)
Business model: monthly subscription	
"I would be willing to pay a regular monthly subscription." *"Ich wäre bereit, regelmäßig ein monatliches Abo zu zahlen."*	0.06
Business model: shared monthly subscription	

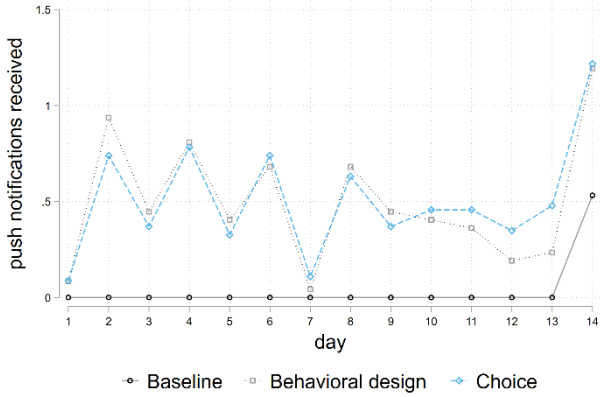
"I would be willing to pay a regular monthly subscription that I could share with family and friends. and friends." *"Ich wäre bereit, regelmäßig ein monatliches Abo zu zahlen, das ich mit Familie und Freunden teilen kann."*	0.10
Business model: per article	
"I would be willing to pay per article read." *"Ich wäre bereit, pro gelesenem Artikel zu bezahlen."*	0.03
Business model: donation	
"I would be willing to make a donation to the provider of the otherwise free app from time to time." *"Ich wäre bereit, dem Anbieter der sonst kostenlosen App ab und zu eine Spende zukommen zu lassen."*	0.15
Business model: sell data to third parties	
"I would use the app free of charge and in exchange I would agree that my usage data can be sold to third parties." *"Ich würde die App kostenlos nutzen und im Gegenzug wäre ich damit einverstanden, dass meine Nutzungsdaten an Dritte verkauft werden können."*	0.19
Business model: use data for market research	
"I would use the app free of charge and in return I would be willing to occasionally participate in market research and behavioral studies within a regulated framework." *"Ich würde die App kostenlos nutzen und im Gegenzug wäre ich bereit, ab und zu - im geregelten Rahmen - an Marktforschungs- und Verhaltensstudien teilzunehmen. "*	0.56
Business model: use data by private companies	
"I would use the app free of charge and in return I would be willing to occasionally allow for my usage data to be send to private companies within a controlled framework." *"Ich würde die App kostenlos nutzen und im Gegenzug wäre ich bereit, ab und zu - im kontrollierten Rahmen - meine Nutzungsdaten an private Unternehmen zu senden."*	0.13
Business model: use data for non-profit research	
"I would use the app for free and in return I would be willing to donate my usage data to universities for research purposes." *"Ich würde die App kostenlos nutzen und im Gegenzug wäre ich bereit, meine Nutzungsdaten zu Forschungszwecken an Universitäten zu spenden."*	0.63
Business model: no compensation	
"I would use the app for free and for no compensation." *"Ich würde die App kostenlos und ohne Gegenleistung nutzen."*	0.64
Business model: not use app	
"I would not use the app under any circumstances." *"Ich würde die App unter keinen Umständen nutzen."*	0.04
Business model: other	
"I would use the app on terms other than the above." *"Ich würde die App zu anderen als den obigen Konditionen nutzen."*	0.09

Figure A1: Usage of and exposure to behavioral elements by treatment groups and over time

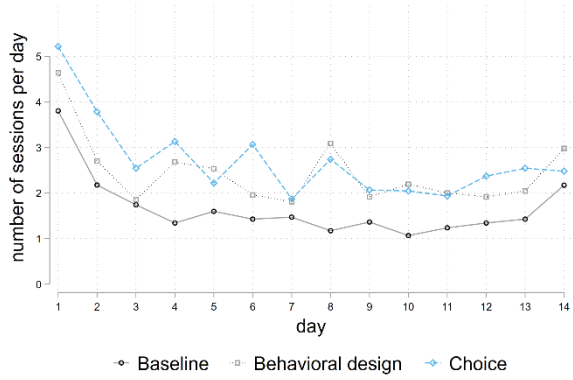
Pull-to-refresh usage



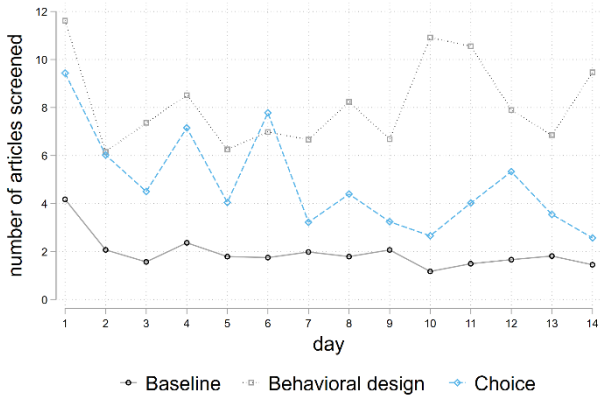
Push notifications received



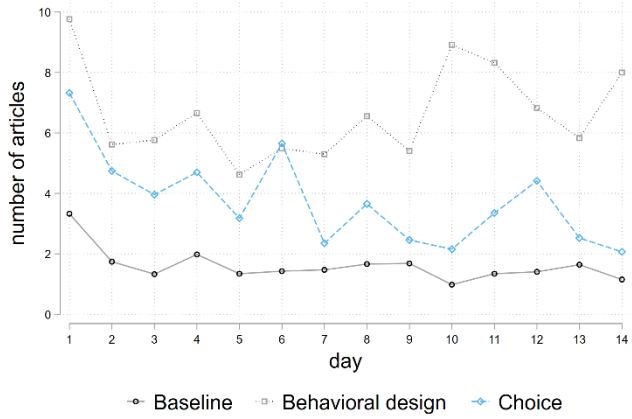
Number of sessions per day



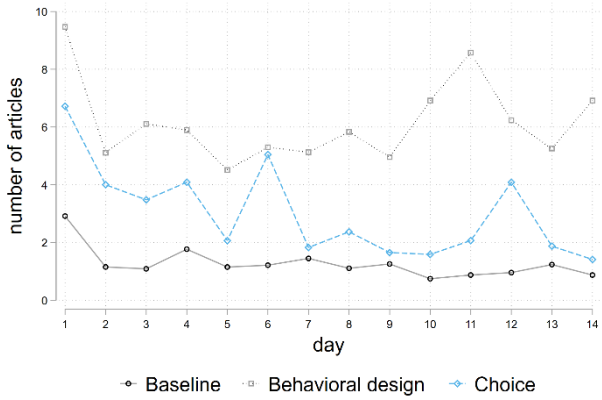
Number of articles screened



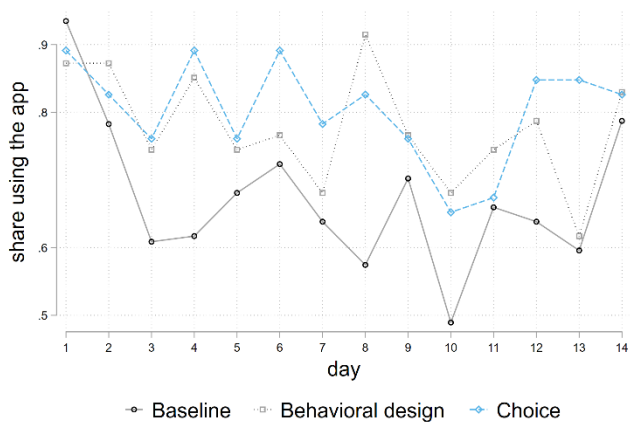
Number of articles read



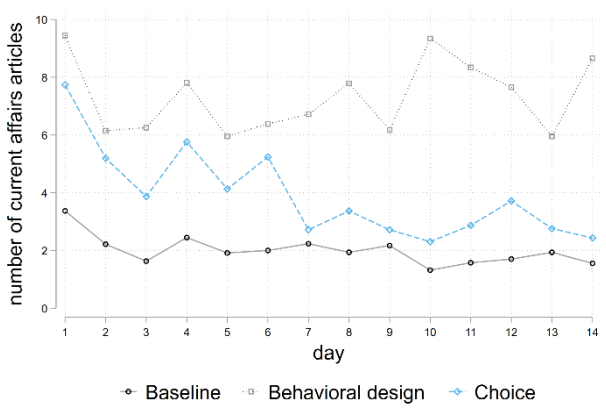
Number of articles rated



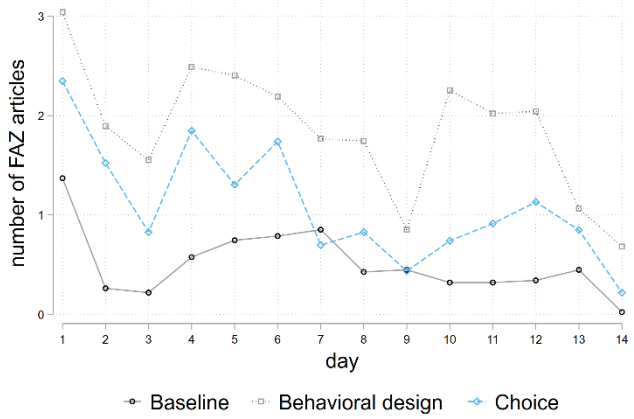
Share of participants using the app



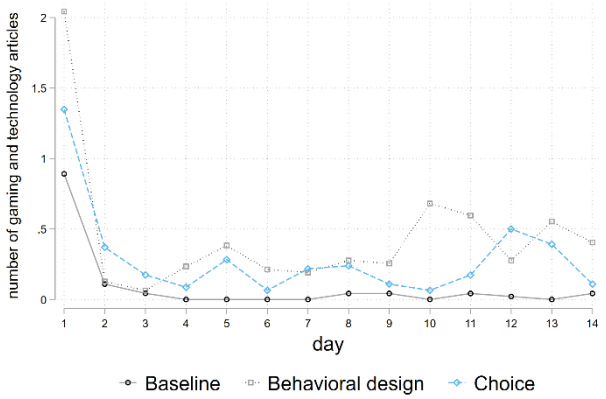
News category: current affairs



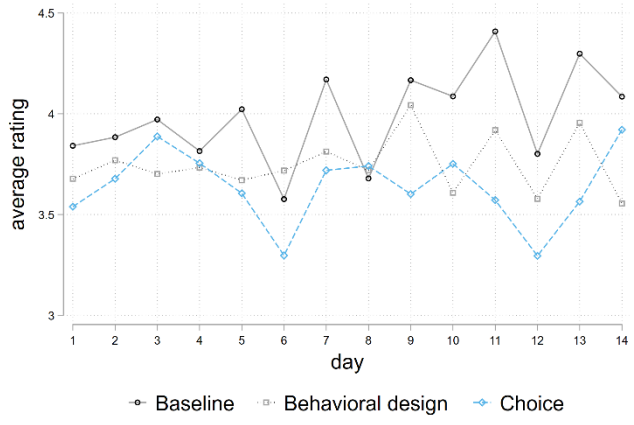
News source: FAZ articles



News category: Gaming and tech articles



Average rating



Average sentiment

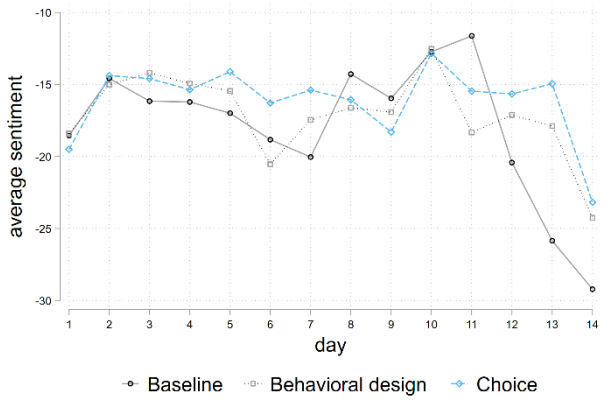
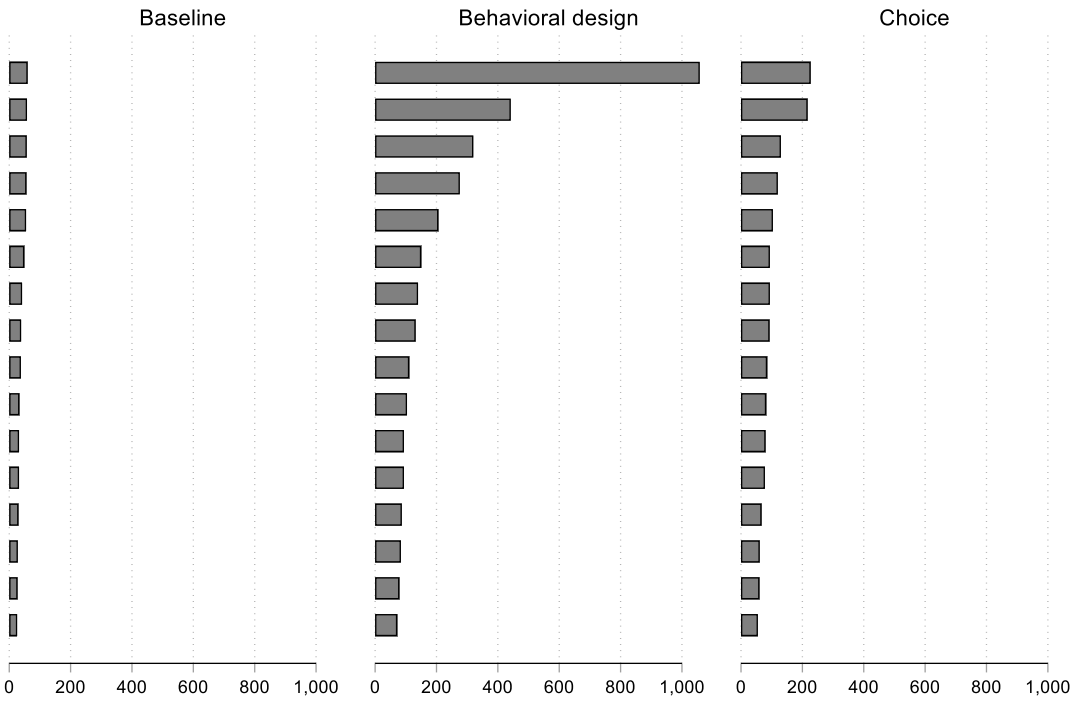
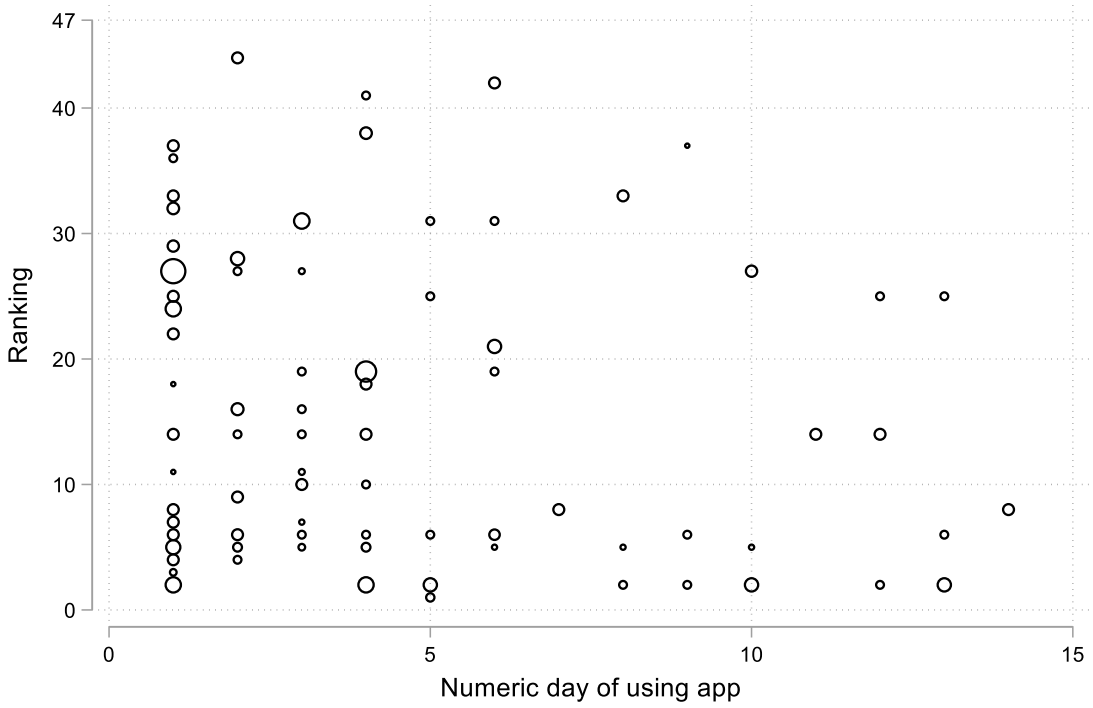


Figure A2: Number of articles read by most active users per treatment group



Notes: For each participant, the articles that were counted as read were added up. The 15 participants who read the most articles in their respective treatment group were ranked from highest to lowest number of articles read.

Figure A3: Usage of filters by study day and user ranking in the *Choice* group



Notes: User ranking according to the total time spent using the app, the size of the circles reflects the number of filter setting changes, conditional on at least one change; the mean is 11.667, the median 7.5 and the max 71.

APPENDIX B: PRELIMINARY STUDIES

Table B1: Ethical considerations by Fogg (2002)

Responsible behavioral design by Fogg (2002)
Ethical considerations related to the behavioral design of information technology root back to the early days of its appearance. Fogg (2002) pioneered a discussion on requirements for responsible behavioral design. The requirements can be summarized as: (1) Designers and app providers shall be aware of the consequences, the potential for backfire, and risks of the applied behavioral design elements and have tested their effects thoroughly. (2) At the same time, consumers shall get informed and educated about the applied behavioral design as well as its risks, such as losing money, privacy, or freedom. (3) And, consumers shall get the chance to interactively make necessary adjustments to protect themselves from unwanted behavioral design.

Table B2: Interviews with software engineers

Interviews with software engineers
Background on design process
For the study app to be as close as possible to real-world apps, we aimed to mimic app designers' design processes. How do app designers prepare their design process? Based on the technology acceptance model, perceived usefulness, perceived ease-of-use, and user acceptance are the cornerstones of successful app design (Davis, 1989). But which standards do app designers actually follow? And in which reflective processes do they engage? To answer these questions, literature grants some insights related to the privacy and security behaviours of app designers (Balebako et al. 2014). When making design decisions, they also keep in mind business interests and company directives. Moreover, they lack awareness of the privacy and security risks when using third party software (Balebako et al. 2014). However, there is a gap in the literature related to the risk awareness of behavioural design in apps, so we interviewed software engineers in a preliminary about this topic.
Description
In a semi-structured interview study in Germany in October and November 2020, twelve software engineers of various backgrounds and genders talked about their values and attitudes and how these match with the expectations of their customers and contractors, and with established business models.

We recruited the software engineers through various channels, such as the Institute for IT Security and Software Development, agencies for software development, centres, associations, educators and advisors of the IT sector, internet forums and platforms. Seven out of the twelve software engineers were active in fields that also require IT consulting and IT project management. Other reported fields of activities covered programming, admin services, hosting, web design, user experience design, game development, insurances, IT security, data mining, and IoT. Half of the participants worked at big companies with more than 250 employees and half of them worked at small and midsized companies.

Deductive coding was used for the qualitative analysis of exploratory questions. The Straussian grounded theory was applied for coding and interpretation of the open-ended questions.

Aim of the interviews

The interview questions dealt with topics around risks of automated decision-making and ubiquitously applied behavioral design, such as risks of discrimination and dependencies on technologies. Resulting needs for regulation, possible measures for consumer protection, and alternative incentives and business models were discussed with the software engineers too. Moreover, we asked them about responsible application of behavioural design and how they assess the prevalence of less ethical practices in their sector. Out of this information, we were able to construct a theory explaining the role of software engineers in influencing whether behavioural design will be applied responsibly or not.

Results

Trustworthiness and the quality of responsible applications was the most important value that software engineers named for success.

Software engineers reported about their critical considerations during the design process, experiencing a value conflict between business goals and user welfare, especially if the business addresses third parties, rather than directly the users. *"If nothing is paid, the user is not the client but the traded goods."* - as one of the interviewees said. However, such concerns are rarely addressed by contractors. Thus, if software engineers do not like a project on moral grounds, they can decide to leave. But in return, they observe a systematic moral erosion in the sector, from one generation of developers to the next, as competitive pressure grows. As one interviewee put it: *"Even if I leave the project, there will be the next one doing the job."*

App designers are in the position of being authentic sources of best advice for users, because they know the design details and during the design process they may consult their inner moral compass and flag questionable practices.

Software engineers reported to us that they are often pressured to imitate successful trends. For example, if the like-button has been a success, it gets copied in other apps too. Following trendsetters is similar to using third party software in the sense that app designers do not question what they copy, or they ignore it for the above-mentioned moral erosion phenomenon. As they report, often they are not aware of the underlying psychological mechanisms and behavioral

consequences of the behavioral design they apply. They are not aware of any risks or backfiring for users. Consequently, they do not have any standardised process in place to control for such risks.
Action taken
To fill this gap, we reviewed the literature, collected well-known behavioral taxonomies, extracted their behavioral design elements, and ordered them into a new, unifying systematics of behavioral design elements (for more details, see Table C1 in Appendix C).

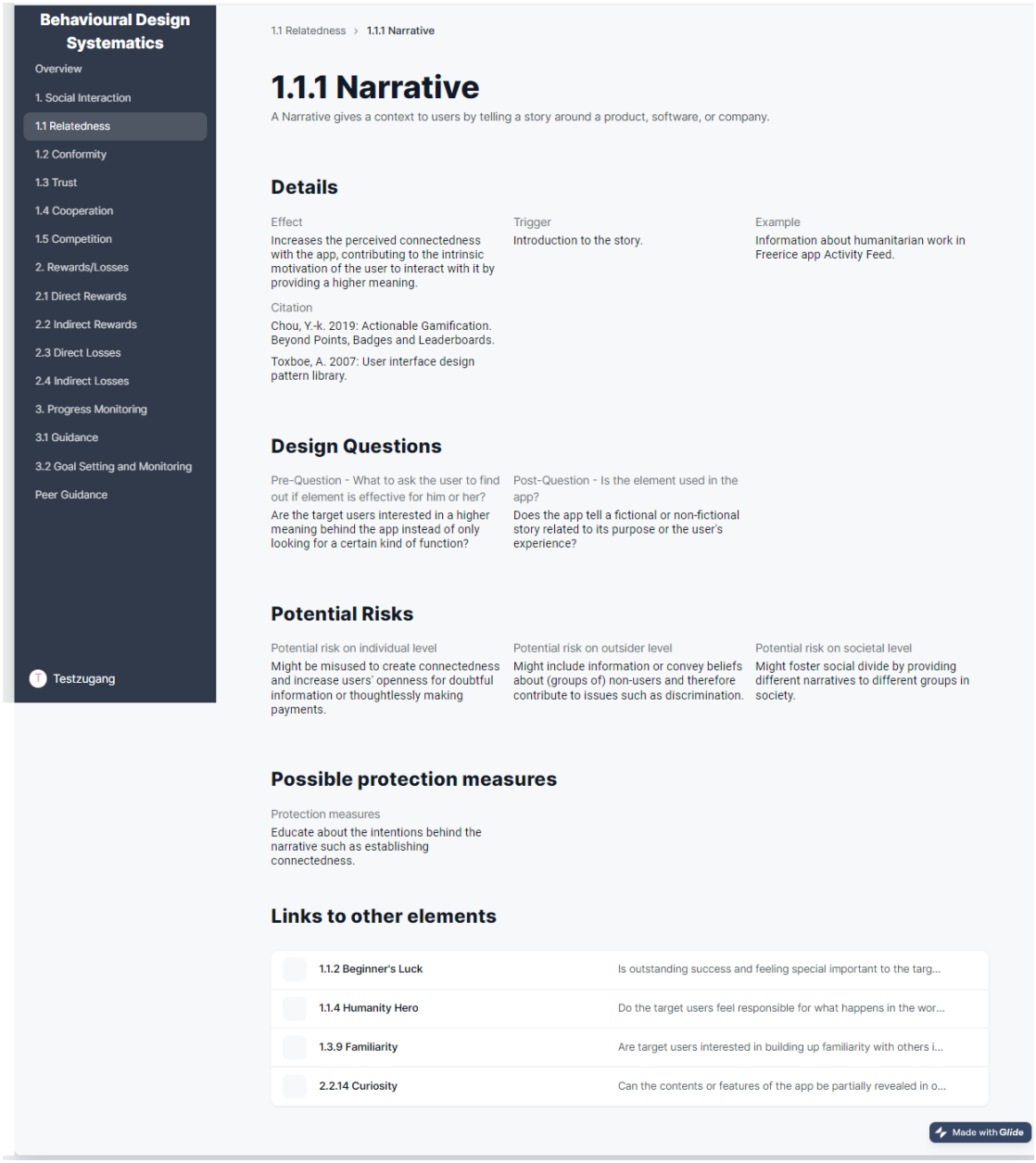
Table B3: Pre-study survey on user profile

User profile
Description
Prior to the field experiment, from February to March 2021, we conducted an independent online study survey among 77 students, employees, and professors at a German public university. This sample was different from our study sample, because the distance between the two universities (pre-study and main study) is about 400 km.
Aim of the pre-study survey
We asked participants about their reading behavior and needs, if reading takes place using a news app.
Results
Unsurprisingly, participants turned out to be a demanding and conscious audience that reads several times a day and many times in between. We found that an app design that motivates readers to read more frequently, while not sending more than 1-2 push notifications a day, has good chances to increase reading frequency and by that the total reading time. Social aspects, such as comparing reading performance with other readers or peer rating of news, manifold and transparent options to filter news, and self-selection of news content are all features that are welcomed by the participants of the pre-study. Distracting components, such as gamification, ads, tabloid journalism, and filter bubbles should be avoided, according to the participants. These insights entered into the design of the protection measures, e.g. in the form of raising awareness about gamification and allowing adjustments to filtering outcomes.

APPENDIX C: SHOWCASING DESIGN SYSTEMATICS

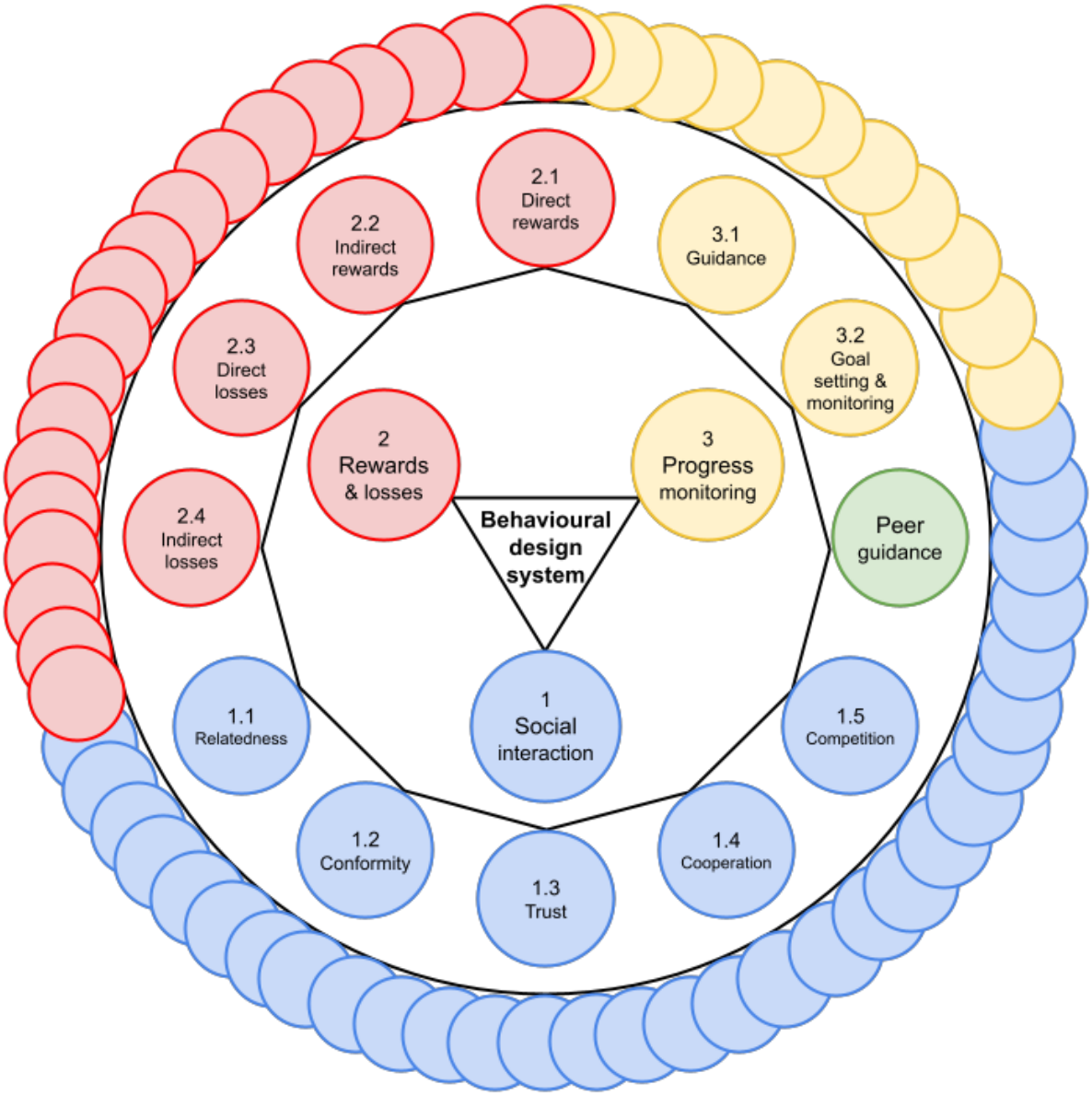
Table C1: Systematics of behavioral design elements

Systematics of behavioral design elements
Description
Digital media applications implement rich arsenals of behavioural design elements when creating inescapable decision-making environments for users. These elements have been categorized by many authors before, based on various perspectives.
Chou (2019) put focus on gamification that intends to increase entertainment and attractiveness of applications, while binding users through habit-formation.
Michie et al. (2013) focus on benevolent behavioural change, for example to promote health prevention. The authors invented the so-called Behavioural Change Wheel. This tool provides an extensive, consensually agreed and hierarchically structured taxonomy of 93 behaviour change techniques used in behaviour change interventions.
Dolan et al. (2012) created a tool called Mindspace, which focuses on the nine most robust effects that influence human behaviour in mostly automatic, rather than deliberate, ways.
Oinas-Kukkonen & Harjumaa (2008) put the focus on the systematic and evaluative design process and proposed the Persuasive Systems Design framework.
Fogg (2002) laid down what outlines persuasive technology and how it changes what humans think and do. Along this line, Fogg established the Fogg Behaviour Model, which shows that three factors must converge at the same moment for a behaviour to occur: motivation, ability, and a prompt. Hence, an optimal timing and optimal interaction of these factors determines the design. In the Fogg Behaviour Grid, he distinguishes various behaviour intensities (new, familiar, increased intensity, decreased intensity, stopped behaviour) and time intervals (one-time, longer duration, permanent change).
Numerous other approaches exist for systematizing behavioural design elements. For example, Toxboe (2007) maintains a user interface design pattern library. Brignull (2010) defined deceptive design and established the first list of dark patterns. Mathur et al. (2021) provides one of the first conceptual foundations that helps to judge what makes a dark pattern dark. Based on a systematic literature review, Mirsch et al. (2017) compiled a comprehensive overview of relevant psychological effects and exemplary nudges, especially in digital environments. In the systematics, these vast insights get united.
Behavioral design elements



The systematics addresses the need for a catalogue that integrates comparable details of each commonly known behavioral design element. Such elements are, for example, praise (i.e. positive feedback for performing a desired action), leaderboards (i.e. ranking that allows users to compare their performance based on a certain goal that is usually connected to the desired action), or fear of missing out (i.e. concerns of missing some interesting or exciting event or feature). The details of each element include name, description, background knowledge, citation of sources for further reading, potential effects and triggers, an exemplary application, reflective design questions easing the reflection on the suitability of applying the element, specification of potential risks, examples for protection measures, and links to other elements. The latter help to achieve concertedness within an app.

Systematics



The systematics currently contains 98 behavioural design elements. These elements are ordered into three overarching categories based on motivational drives: social interaction, rewarding, and progress monitoring. Based on similarities in the behavioral design elements, the overarching categories have been subdivided into nether categories, to further ease the designer's search for certain elements. Social interaction contains relatedness (10 elements), conformity (9 elements), trust (10 elements), cooperation (4 elements), and competition (4 elements). Rewards and losses contain direct rewards (12 elements), indirect rewards (18 elements), direct losses (4 elements), and indirect losses (7 elements). Progress monitoring contains guidance (12 elements) and goal setting and monitoring (7 elements). Peer guidance both fits to social interaction and guidance in progress

monitoring. The categories enable quick orientation in the complex network of the behavioural design elements.

The outer ring represents the 98 single behavioural design elements, which were categorized into the three overarching (inner ring) and nether (middle ring) categories. Each overarching category is coloured differently.

Weblink

(Hidden for anonymity during the review process.)
The systematics is featured as a handy tool in the form of a web-based app, which is a novel contribution. The tool may be useful for appliers of behavioural design aiming to deepen their understanding, users aiming for more transparency, and future research extending on this work.

Designing for transparency

Based on the systematics and its critical highlighting of questionable behavioral design elements, we identified the most questionable elements in our app and designed adjustable design features for those. For example, the algorithmic news filtering that personalizes the news feed got a local button that allowed (1) to monitor the automated filter settings, (2) to adjust those manually, or (3) to turn them off. When designing the transparency features, we applied the technology acceptance model (Davis, 1989). For example, the local button to adjust the filter is convenient and user-friendly. Similarly, the informed consent to behavioral design is written in easy language, accompanied by icons easing the orientation, and showing the otherwise also locally adjustable features (i.e. users can make the adjustments locally or globally in the app).

Thus, the main design principles of the transparent app version are: (1) enable alternative choices, (2) make it easy to make decisions, (3) avoid overloaded design solutions by focussing, (4) break down complex tasks and situations into smaller steps (e.g. using local buttons that are either intuitive or accompanied by short explanations), (5) integrate the joy of exploring with the step-by-step training and mastering of digital media competency related to behavioral design.

Benefits & Contribution

We used the systematics for the design of our study app. It helped us to select and concert behavioral design elements to a complex holistic app design, so that the elements function harmoniously together as part of the whole app. As a proof of concept, we showed how the tool facilitates the design process and how protection measures might be implemented into apps. The new tool is an important contribution, because it makes responsible design easy and easily scalable for millions of apps existing.

For us using the systematics was crucial, because our research budget allowed for only one shot, whereas app designers usually test each design element via A/B testing (i.e. comparing two versions of an app against each other to determine which one performs better) among their vast user pool. Hence, using the systematics opens up a possibility to save app development resources.

For more details on which elements we selected, see Table C2 in Appendix C.

Table C2: Complete list of the applied behavioral design elements and protection measures

Behavioral design elements & protection measures		
Element	**In the app**	**Explanation**
Onboarding		
Appearance **Similarity** Fogg 2002 Durantini et al. 2007 Oinas-Kukkonen & Harjumaa 2008 Dolan et al. 2012 **Surface Credibility** Oinas-Kukkonen & Harjumaa 2008	Whooop	The appearance of the app (e.g. visual cues and color convention) is designed to be coherent, neat, recognizable, and appealing in order to catch attention. The name of the app, Whooop with three strongly pronounced 'o'-s, reminds of a sudden exclamation calling for attention or expressing amazement. The logo symbolizes the wondering face of a stylized sloth forming an 'o'. It appears also in the lettering in a playful manner. The sloth was chosen as a symbol for the 'dull user', who deliberately chose to be entertained. The emphasis is on the sloth's choice and its ability to adapt to its designed environment. The choice should be deliberate, rather than invisibly governed by behavioral design. With this logo we aim to raise awareness even beyond the study, throughout the publication and dissemination work. **Similarity** is applied by reminding the amazed reader of herself and the feeling accompanying the discovery of interesting news. Thus users feel more connected to the app and are more open to being influenced by it when they recognize something similar to their own character or (social) environment. **Surface Credibility** means that the app looks and feels serious and competent by its neat appearance and accurate language. This element can lower the barrier for users to download and use the app. **Present in:** Baseline Behavioral Design Choice
Launch screen		

Narrative Toxboe 2007 Chou 2019 **Curiosity** Loewenstein 1994 Toxboe 2007 Jirout 2020 **Tunneling** Fogg 2002 Toxboe 2007 Oinas-Kukkonen & Harjumaa 2008 **Priming** Lashley 1951 Toxboe 2007 Mirsch et al. 2017		The Launch screen is onboarding participants. The first sentence says "Successful people spend an above-average amount of time reading news". Such and similar statements form a ***Narrative*** about the usefulness of reading news. This framing narrative is continued in the tutorial by emphasizing benefits of the app rather than simply introducing features. The narrative in turn contributes to ***Priming*** participants to make them eager to read as much as possible. Additionally, it marks the beginning of the tutorial process that follows a ***Tunnelling*** approach, ensuring that participants understand all relevant functions and set up their account properly (see following screens). ***Curiosity*** is instilled by giving a first insight of Whooop's functions without revealing everything at once. **Present in:** Baseline Behavioral Design Choice 
Tutorial 1 **Narrative** Toxboe 2007 Chou 2019 **Curiosity** Loewenstein 1994 Toxboe 2007 Jirout 2020 **Tunneling** Fogg 2002 Toxboe 2007 Oinas-Kukkonen & Harjumaa 2008 **Priming** Lashley 1951 Toxboe 2007 Mirsch et al. 2017 **Progress Bars** Toxboe 2007 Chou 2019		The first Tutorial screen gives more details on the core function of the app, the news feed, and continues the ***Tunneling*** process and the approach of keeping up ***Curiosity***. ***Priming*** is strengthened by the message "Stay up to date and save valuable time by seeing all relevant news at a glance". The ***Progress Bar*** at the bottom of the screen provides orientation along the tutorial process, keeping participants motivated to complete the tutorial. **Present in:** Baseline Behavioral Design Choice 
Tutorial 2		

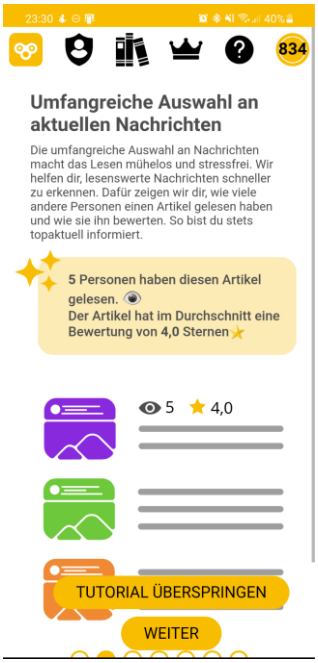
Social Facilitation
Guerin 1993
Oinas-Kukkonen & Harjumaa 2008

Curiosity
Loewenstein 1994
Toxboe 2007
Jirout 2020

Tunneling
Fogg 2002
Toxboe 2007
Oinas-Kukkonen & Harjumaa 2008

Priming
Lashley 1951
Toxboe 2007
Mirsch et al. 2017

Progress Bars Toxboe 2007
Chou 2019



The second Tutorial screen introduces the views and rating functions. Here, **Social Facilitation** is added by displaying the behavior of other participants (reading and rating articles), making it easier for the individual to adopt the behavior.

(For the other elements see the screen before.)

Present in:

Baseline
Behavioral Design
Choice

Tutorial 3

Social Facilitation
Guerin 1993
Oinas-Kukkonen & Harjumaa 2008

Peer Rating
Toxboe 2007

Curiosity
Loewenstein 1994
Toxboe 2007
Jirout 2020

Tunneling
Fogg 2002
Toxboe 2007
Oinas-Kukkonen & Harjumaa 2008

Priming
Lashley 1951
Toxboe 2007
Mirsch et al. 2017

Progress Bars Toxboe 2007
Chou 2019

Here, in combination with other elements, **Priming** focuses on making users contribute to another element: **Peer Rating**, which increases trust and perceived connection with other users.

(For the other elements see the screens before.)

Present in:

Baseline
Behavioral Design
Choice

Tutorial 4

Status Points
Schell 2008
Chou 2019

Reward Oinas-Kukkonen & Harjumaa 2008
Schell 2008
Michie et al. 2013

Curiosity
Loewenstein 1994
Toxboe 2007
Jirout 2020

Tunneling
Fogg 2002
Toxboe 2007
Oinas-Kukkonen & Harjumaa 2008

Priming
Lashley 1951
Toxboe 2007
Mirsch et al. 2017

Progress Bars Toxboe 2007
Chou 2019

For each article that a participant has read, he or she is awarded three points as a **Reward**. This kind of **Status Points** allows participants to track their progress and can motivate them to consistently perform the rewarded action (in this case reading news on a regular basis).

(For the other elements see the screens before.)

Present in:

Behavioral Design
Choice

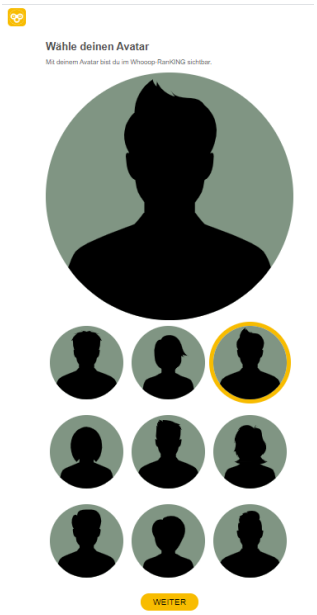
Avatar

Anonymity

Personalization Oinas-Kukkonen & Harjumaa 2008
Schell 2008

Need for Closure
Barlow 1981
Toxboe 2007
Schell 2008

Tunneling
Fogg 2002
Toxboe 2007
Oinas-Kukkonen & Harjumaa 2008



Users who finished the tutorial stage and are then prompted to complete the onboarding process might feel a **Need for Closure**, motivating them to stay inside the app until the process of setting up and personalizing the app is finished. Choosing an avatar allows for **Personalization** without forcing the user to give up their **Anonymity**, keeping the barrier for using the app as low as possible. (Besides this effect of **Anonymity**, it was a standard procedure for all study participants to preserve their anonymity within the lab's framework.) In turn, having made a step towards a personalized experience can increase the willingness to complete the following steps of accepting terms and conditions as well as allowing the app to use any personal and behavioral data.

Present in:

Behavioral Design
Choice

Protection measures during onboarding & later globally accessible

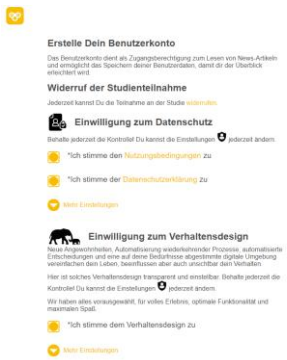
In-app consent

Reduction
Fogg 2002
Toxboe 2007
Oinas-Kukkonen & Harjumaa 2008

Tunneling
Fogg 2002
Toxboe 2007
Oinas-Kukkonen & Harjumaa 2008

Serial Positioning
Ebbinghaus 1913
Toxboe 2007

Defaults
Cronqvist & Thaler 2004
Toxboe 2007



Elements for the in-app consent are placed in a certain order by applying **Serial Positioning**, making necessary consents for data protection the first priority by presenting them first. More complex and optional information and adjustments are presented at the bottom of the screen. **Reduction** makes it easy to complete the process quickly without having to open the fold-down menu if the participant is not interested in details. **Defaults** additionally save time and effort.

A special novelty as part of the in-app consent is the ***"Consent to Behavioral Design"*** in the *Choice Design* version. (The other app versions have only a "Consent to Data Protection".) As an educational protection measure, this measure aims at raising awareness of mechanisms of behavioral design.

Present in:

Baseline
Behavioral Design
Choice *(has extended version)*

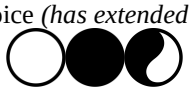
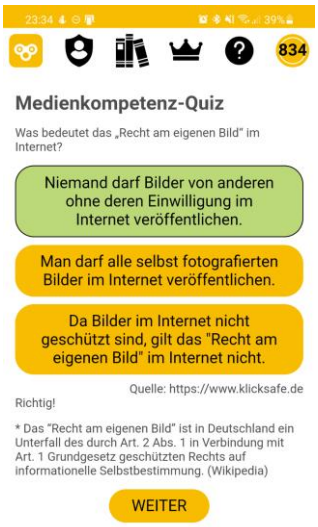
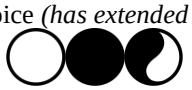
Fold-down menu of in-app consent

Trustworthiness Oinas-Kukkonen & Harjumaa 2008

Reduction
Fogg 2002

If the participant opens the fold-down menu (**Reduction**, see previous row), further options are visible that are turned on via **Defaults**. If participants are interested, they can switch those on or off, allowing them to control their settings and establishing **Trustworthiness**.

Present in:

Toxboe 2007 Oinas-Kukkonen & Harjumaa 2008 **Tunneling** Fogg 2002 Toxboe 2007 Oinas-Kukkonen & Harjumaa 2008 **Serial Positioning** Ebbinghaus 1913 Toxboe 2007 **Defaults** Cronqvist & Thaler 2004 Toxboe 2007		Baseline Behavioral Design Choice *(has extended version)* 
Media competency quiz **Trustworthiness** Oinas-Kukkonen & Harjumaa 2008 **Halo Effect** Thorndike 1920 Kahneman 2011		During the in-app consent, participants get the option to do a quiz that helps them assess their media competency, providing guidance when deciding if they would like to skim the content and give their consent quickly or if they should read it more carefully. This can support **Trustworthiness** and, in combination with other protection measures, possibly contribute to a kind of **Halo Effect**, making the app appear especially trustworthy and concerned with the participants' data sovereignty (and awareness of and protection from behavioral design in the *Choice Design* version) throughout the whole user experience. **Present in:** Baseline Behavioral Design Choice *(has extended version)* 
Core function		
News feed		Similar to the algorithms that big companies like Facebook apply in their apps' news feeds (although much simpler),

Trustworthiness Oinas-Kukkonen & Harjumaa 2008

Tailoring
Fogg 2002
Toxboe 2007
Oinas-Kukkonen & Harjumaa 2008

Personalization Oinas-Kukkonen & Harjumaa 2008
Schell 2008

Pull-to-Refresh Toxboe 2007
Harris 2016

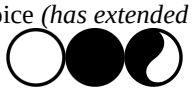
Infinite Scroll
Toxboe 2007
Harris 2016

Defaults
Cronqvist & Thaler 2004
Toxboe 2007

Tailoring is used to make the news feed adapt to participants' interests, showing them more articles that are similar to their preferred reading choices, which can in turn motivate them to read even more articles. The option to filter news by certain attributes (e.g. top news, see yellow article in the screenshot), is a form of **Personalization**, making the participant feel in control. **Pull-to-Refresh** is applied as a mechanism for loading new content into the news feed, contributing to the development of a potential habit - if new content appears, which works as a kind of reward for performing the gesture. **Infinite Scroll** is aimed at keeping participants engaged with the app while scrolling through a seemingly endless stream of news. Presenting the (reliable) sources of the news contributes to **Trustworthiness**. By **Default**, the option is active, inviting the participant to use it and see what is behind.

Present in:

Baseline *(has simplest version with 'Trustworthiness' and 'Infinite Scroll')*
Behavioral Design
Choice *(has extended version)*



Article

Social Facilitation
Guerin 1993
Fogg 2002
Oinas-Kukkonen & Harjumaa 2008

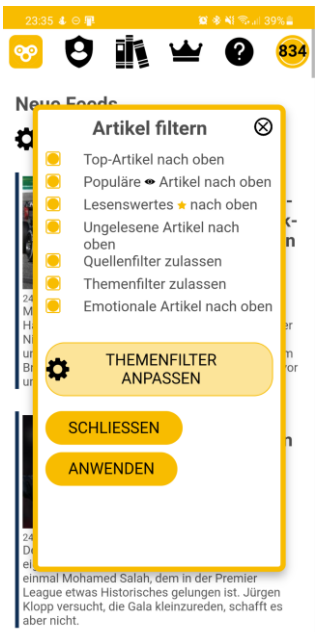
Trustworthiness Oinas-Kukkonen & Harjumaa 2008

Peer Rating
Toxboe 2007

When the participant opens an article and scrolls down, the **Peer Rating** function becomes visible. If other participants have already rated the article, the average rating will be displayed here, providing **Social Facilitation**. **Trustworthiness** is created again by providing the source of the article. The **Isolation Effect** makes it easy to find the "Back" button at the end of the page when reading is finished, bringing the participant back to the news feed where more articles are waiting.

Present in:

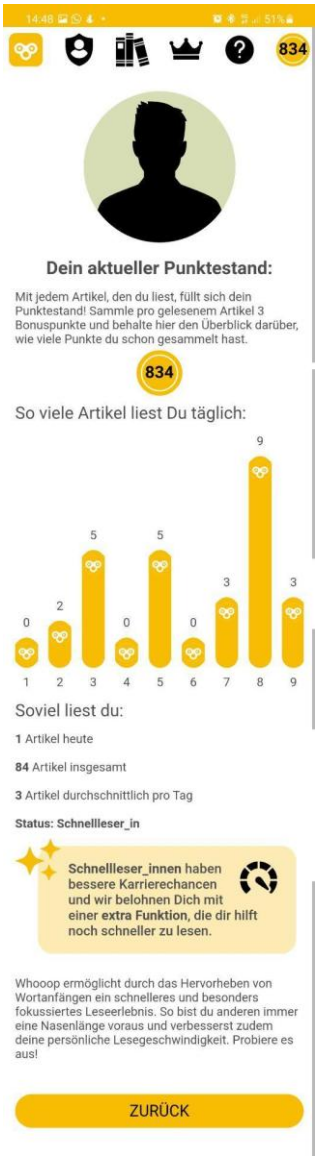
Isolation Effect Von Restorff 1933 Toxboe 2007 Chou 2019		Baseline Behavioral Design Choice 
Protection measures for the core function		
Filter option **Isolation Effect** Von Restorff 1933 Toxboe 2007 Chou 2019 **Personalization** Oinas-Kukkonen & Harjumaa 2008 Schell 2008		In the *Choice Design* version, the filter can be turned off locally using a toggle switch. The toggle switch is an outstanding element with an **Isolation Effect** guiding the user towards the protection measure. Moreover, the filtering options can be adjusted using local settings, like a kind of **Personalization**. **Defaults** show how filtering works and invite the participant to explore the different possibilities. **Present in:**

Defaults Cronqvist & Thaler 2004 Toxboe 2007		Choice
Correction of algorithmic results **Self-Monitoring** Oinas-Kukkonen & Harjumaa 2008 Michie et al. 2013		In the *Choice Design* version, the result of algorithmic filtering is shown transparently, so that **Self-Monitoring** and self-administered adjustments become possible. **Present in:** Choice
Rewards & losses		
Achievements **Achievements** Toxboe 2007 Chou 2019 **Reward** Oinas-Kukkonen & Harjumaa 2008		Different **Achievements** can be unlocked by consistently reading articles (e.g. the "Beginner" achievement, a **Reward** for reading the first article). This **Collection Set** motivates participants to set increasingly demanding goals and spend more time reading articles in order to collect as many achievements as possible. **Present in:**

Schell 2008 Michie et al. 2013 **Collection Sets** Chou 2019		Behavioral Design Choice
Achievement message **Beginner's Luck** Chou 2019 **Achievements** Toxboe 2007 Chou 2019 **Praise** Toxboe 2007 Oinas-Kukkonen & Harjumaa 2008 Schell 2008		Participants are informed about having unlocked an **Achievement** through an in-app message, stating in this case "Whooop, you read an article!" for the "Beginner" achievement. This achievement is unlocked at the very beginning of the user experience as an instance of **Beginner's Luck**, making the participant feeling successful from the beginning. Generally, the messages aim at encouraging participants to read more in order to unlock more achievements by relying on the element **Praise**. Again, this works as a **Reward** for performing the desired action of reading articles and spending more time with the app. **Present in:**

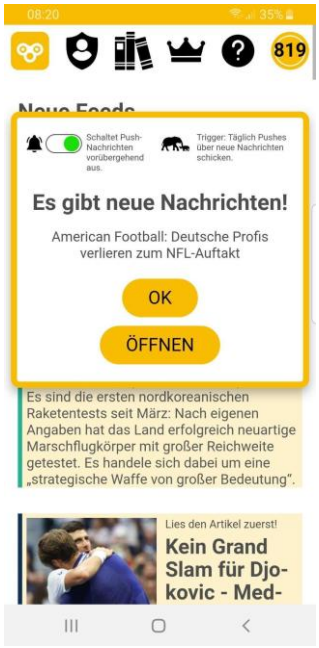
Reward Oinas-Kukkonen & Harjumaa 2008 Schell 2008 Michie et al. 2013		Behavioral Design Choice
Social interaction		
Reader RanKING **Social Facilitation** Guerin 1993 Fogg 2002 Oinas-Kukkonen & Harjumaa 2008 **Conformity Anchor** Chou 2019 **Social Comparison** Oinas-Kukkonen & Harjumaa 2008 Michie et al. 2013 **Leaderboards** Toxboe 2007 Chou 2019 **Recognition** Oinas-Kukkonen & Harjumaa 2008		Participants can compare their bonus points for the day (daily ranking) and in total (overall ranking) with the points that others achieved so far in a **Leaderboard**. On the one hand, this contributes to a general **Social Facilitation**, on the other hand, the dynamics of **Social Comparison** (motivation through comparing the own performance to others') and **Conformity Anchor** (wish to make the own behavior fit into social norms, in this case reading much) push participants to read more if others do so, too. Dividing rankings into an overall and a daily ranking gives all users the opportunity to rank high and receive **Recognition** from others. **Present in:** Behavioral Design Choice
Confetti for best reader **Praise**		The best reader gets an additional motivational boost in the ranking section in the form of visual **Praise**.

Toxboe 2007 Schell 2008 Oinas-Kukkonen & Harjumaa 2008		**Present in:** Behavioral Design Choice
Progress monitoring		
Personal record **Elitism** Chou 2019 **Status Points** Schell 2008 Chou 2019 **Easter Eggs/Sudden Rewards**		At any time, participants can open their personal record area, **Self-Monitoring** their progress and reading habits as well as their **Status Points**. Additionally, for participants belonging to the half of the sample that read more words per minute on average, an extra function supporting fast reading was unlocked after two weeks. This serves as an **Easter Egg/Sudden Reward**, aiming at increasing interest in the app and instilling a subtle feeling of **Elitism** for those who now belonged to the group of “Fast Readers”. **Present in:**

Deci & Ryan 1985 Toxboe 2007 Chou 2019 **Self-Monitoring** Snyder 1974 Fogg 2002 Toxboe 2007 Oinas-Kukkonen & Harjumaa 2008 Michie et al. 2013		Behavioral Design Choice
Notifications		
Push notifications **Curiosity** Loewenstein 1994 Toxboe 2007 Jirout 2020 **FOMO** Chou 2019 **Reminders** Toxboe 2007		Visual signaling and vibration triggers a new interaction with the app: opening the new incoming message, relying on both **Curiosity** and **FOMO**, the fear of missing something important. This common type of **Reminders** might lead to a serious checking-habit as a potential risk (Palokangas, 2016). **Present in:**

Oinas-Kukkonen & Harjumaa 2008		Behavioral Design Choice 
Push notifications - Content **Leaderboards** Toxboe 2007 Chou 2019 **Appointment Dynamics** Toxboe 2007 Chou 2019 **Loss Aversion** Kahneman & Tversky 1979 Toxboe 2007 Dolan et al. 2012 Mirsch et al. 2017 **Reminders** Toxboe 2007 Oinas-Kukkonen & Harjumaa 2008 **Suggestions** Oinas-Kukkonen & Harjumaa 2008 **Simulation** Fogg 2002 Fogg 2007 Toxboe 2007 Oinas-Kukkonen & Harjumaa 2008	*You're falling behind in the daily ranking! Read at least 2 more articles to defend your rank!* Today's daily ranking will close in 4 hours.	Several push notifications (**Reminders**) were integrated into the participants' experience, conveying various messages and relying on different mechanisms. This example utilizes the fact that many people feel stronger motivation to act in order to avoid losing something that they already have (their position in the **Leaderboard**) than in order to gain something, a phenomenon known as **Loss Aversion**. The **Suggestion** of reading 2 more articles serves as a call to action while the **Appointment Dynamics** of the closing ranking can contribute to a feeling of urgency and support tendencies to increase reading efforts at a certain time of the day. The projection of the outcome if the participant does not take action relies on the **Simulation** element. **Present in:** Behavioral Design Choice 
Push notifications - Content **Leaderboards** Toxboe 2007 Chou 2019 **Reward** Oinas-Kukkonen & Harjumaa 2008	*You have almost reached the top 10!* Read 2 more articles to make it into the total ranking.	This push notification combines **Leaderboards**, **Suggestions, Reminders** and **Simulation** as well, but in contrast to the previous example aims at the participant's motivation to reach the more prestigious total ranking (**Reward**), promising **Recognition**. This can be supported by being framed as an **Evanescent Opportunity** (if the participant does not take action immediately, the chance to get into the total ranking might vanish). **Present in:**

Schell 2008 Michie et al. 2013 **Recognition** Oinas-Kukkonen & Harjumaa 2008 **Evanescent Opportunities** Chou 2019 **Suggestions** Oinas-Kukkonen & Harjumaa 2008 **Reminders** Toxboe 2007 Oinas-Kukkonen & Harjumaa 2008 **Simulation** Fogg 2002 Fogg 2007 Toxboe 2007 Oinas-Kukkonen & Harjumaa, 2008		Behavioral Design Choice 
Push notifications - Content **Leaderboards** Toxboe 2007 Chou 2019 **Praise** Toxboe 2007 Oinas-Kukkonen & Harjumaa 2008 Schell 2008 **Reminders** Toxboe 2007 Oinas-Kukkonen & Harjumaa 2008 **Feedback on Performance** Michie et al. 2013	*Great! You moved up in the ranking!* You improved by 5 ranks today. Keep it up!	In this example, the participant is encouraged through **Praise** and a positive **Feedback on Performance**, in order to reinforce the active usage of the app. **Present in:** Behavioral Design Choice 
Push notifications - Content **Status Points** Schell 2008	*Keep it up!* Read 3 more articles to beat your personal high score.	Here, the participant's own **Status Points** are used as a motivation to continue reading.

Chou 2019 **Suggestions** Oinas-Kukkonen & Harjumaa 2008 **Reminders** Toxboe 2007 Oinas-Kukkonen & Harjumaa 2008 **Self-Monitoring** Oinas-Kukkonen & Harjumaa 2008 Michie et al. 2013		**Present in:** Behavioral Design Choice
Protection measures for notifications		
Push notification with protection measures **Reminders** Toxboe 2007 Oinas-Kukkonen & Harjumaa 2008 **Isolation Effect** Von Restorff 1933 Toxboe 2007 Chou 2019 **Personalization** Oinas-Kukkonen & Harjumaa 2008 Schell 2008 **Defaults** Cronqvist & Thaler 2004 Toxboe 2007		In the *Choice Design* version, push notifications include local toggle switches for turning them off. The toggle switches are part of the **Reminder** function. At the same time, toggle switches are outstanding thanks to their **Isolation Effect**, catching attention. The local settings are complemented by global settings under the main menu. Global settings allow turning on push notifications, set a time frame for pausing push notifications, and adjusting **Default** triggers making use of **Personalization** for push notifications and persuasion and habit-formation in the broader sense. *Push notifications* have a key role in developing habits. *Triggers* are circumstances under which a push notification is sent. **Present in:** Choice
Screen time limit **Reminders** Toxboe 2007 Oinas-Kukkonen & Harjumaa 2008 **Isolation Effect** Von Restorff 1933		The screen time limit utilizes **Personalization** and **Defaults**. Via **Reminders** with an **Isolation Effect** a friction in the app is created, which makes the concern with the screen time limit an important task at the present moment. The **Reminder** prompts to pause the app usage, it communicates how long the app has been used in the given session and that "It's time to take a rest for the eyesight." - putting the focus of motivation on a health aspect.

Toxboe 2007 Chou 2019 **Personalization** Oinas-Kukkonen & Harjumaa 2008 Schell 2008 **Defaults** Cronqvist & Thaler 2004 Toxboe 2007		**Present in:** Choice

Notes: The table contains the name of the respective element, the citation of the source, an example of how the element was built into the study app, and an explanation of its effects. Note that the table mainly cites publications which systematized the given element in a behavioral design context. Where possible, the first source mentioning the concept is cited, too. The table includes behavioral design elements applied in protection measures.

APPENDIX REFERENCES

Balebako, R., Marsh, A., Lin, J., Hong, J. and Cranor, L. (2014). The Privacy and Security Behaviors of Smartphone App Developers. 10.14722/usec.2014.23006.

Barlow, A. R. (1981). Gestalt Therapy and Gestalt Psychology, *The Gestalt Journal*, 4 (2).

Brignull H. (2010). Deceptive design. Available at: https://www.deceptive.design/

Chou Y.-k. (2019). *Actionable Gamification. Beyond Points, Badges and Leaderboards*, Packt Publishing Ltd.

Cronqvist H. and Thaler R. H. (2004). Design Choices in Privatized Social-Security Systems: Learning from the Swedish Experience, *American Economic Review*, 94 (2), 424-28. https://doi.org/10.1257/0002828041301632

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly, 13(3), 319–340. https://doi.org/10.2307/249008

Deci E. and Ryan R. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*, New York, NY: Plenum Press.

Dolan P., Hallsworth M., Halpern D., King D., Metcalfe R., and Vlaev I. (2012). Influencing behaviour: The mindspace way, *Journal of economic psychology*, 33 (1), 264-77. https://doi.org/10.1016/j.joep.2011.10.009

Durantini M. R., Albarracín D., Mitchell A. L., Earl A. N., and Gillette J. C. (2006). Conceptualizing the influence of social agents of behavior change: A meta-analysis of the effectiveness of HIV-prevention interventionists for different groups, *Psychological bulletin*, 132 (2), 212. https://doi.org/10.1037/0033-2909.132.2.212

Ebbinghaus H. (1913). *On memory: A contribution to experimental psychology*, New York: Teachers College.

Fogg B. J. (2002). Persuasive technology: using computers to change what we think and do, *Ubiquity*, 2002, 2. https://doi.org/10.1145/764008.76395

Fogg B. J., Danielson D., and Cuellar G. (2007). Motivating, influencing, and persuading users: An introduction to captology. In *The human-computer interaction handbook* (pp. 159-172). CRC press.

Guerin B. (1993). *Social facilitation*, Cambridge University Press, Editions de la Maison des Sciences de l'Homme. https://doi.org/10.1017/CBO9780511628214

Harris, T. (2016). *How Technology is Hijacking Your Mind — from a Magician and Google Design Ethicist*. https://medium.com/thrive-global/how-technology-hijacks-peoples-minds-from-a-magician-and-google-s-design-ethicist-56d62ef5edf3

Jirout J. J. (2020). Supporting Early Scientific Thinking Through Curiosity, *Frontiers in Psychology*, 11, 1717. https://doi.org/10.3389/fpsyg.2020.01717

Kahneman D. and Tversky A. (1979). Prospect theory: An analysis of decision under risk, *Econometrica*, 47, 263-91. https://doi.org/10.2307/1914185

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Lashley K. S. (1951). The problem of serial order in behavior, In *Cerebral mechanisms in behavior*, ed. Jeffress L. A., New York: Wiley, 112-31.

Loewenstein G. (1994). "he psychology of curiosity: A review and reinterpretation, *Psychological Bulletin*, 116 (1), 75-98. https://doi.org/10.1037/0033-2909.116.1.75

Mathur A., Kshirsagar M. & Mayer J. (2021). What Makes a Dark Pattern... Dark? Design Attributes, Normative Considerations, and Measurement Methods. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 360, 1–18. https://doi.org/10.1145/3411764.3445610

Michie S., Richardson M., Johnston M., Abraham C., Francis J., Hardeman W., Eccles M. P., Cane J., and Wood, C. E. (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions, *Annals of behavioral medicine*, 46 (1), 81-95. doi: 10.1007/s12160-013-9486-6

Mirsch T., Lehrer C., and Jung R. (2017). Digital nudging: Altering user behavior in digital environments, Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), 634-48.

Oinas-Kukkonen, H., & Harjumaa, M. (2008). A systematic framework for designing and evaluating persuasive systems. In: Oinas-Kukkonen, H., Hasle, P., Harjumaa, M., Segerståhl, K., Øhrstrøm, P. (eds) Persuasive Technology. PERSUASIVE 2008. Lecture Notes in Computer Science, vol 5033. doi.org/10.1007/978-3-540-68504-3_15

Palokangas L., and Suomala J. (2017). Nudging Problematic Smartphone Use to a Lower Level, *Cognitive Science*. https://api.semanticscholar.org/CorpusID:13220015

Schell J. (2008). *The Art of Game Design: A book of lenses*. CRC press.

Snyder M. (1974). Self-monitoring of expressive behavior, *Journal of Personality and Social Psychology*, 30 (4), 526-37. https://doi.org/10.1037/h0037039

Thorndike, E.L. (1920). A constant error in psychological rating, *Journal of Applied Psychology*, 4, 25-29. https://doi.org/10.1037/h0071663

Toxboe A. (2007). User interface design pattern library, https://ui-patterns.com/patterns/

Von Restorff H. (1933). Über die Wirkung von Bereichsbildung im Spurenfeld [On the effects of the formation of a structure in the trace field], *Psychologische Forschung*, 18, 299-342.

Discussion Paper of the Research Area Markets and Choice 2026

Research Group **Information, Incentives, Inequality**

Maja Adena, Eleonora Alabrese, Francesco Capozza, Isabelle Leader SP II 2026-601
AI Images, Labels and News Demand

Christina Timko, Maja Adena SP II 2026-602
The Impact of Behavioral Design and Users' Choice on Smartphone App Usage and Willingness to Pay: A Framed Field Experiment