Security by behavioural design: a feasibility study

Steen, T. van; Busser, E. de

Security by Behavioural Design: A Rapid Review

Final report for NCSC-NL

Dr. Tommy van Steen

Dr. Els De Busser

Institute of Security and Global Affairs

Leiden University

The Netherlands

Date: 23 July 2021

Table of Contents

Background

1.1 Security by behavioural design

There are broadly two streams of knowledge that the behavioural sciences can contribute to improving cybersecurity behaviour by end-users: 1) by implementing behavioural science insights in the design of the system (security by design), and 2) by developing effective training for end-users (behavioural change campaigns and intervention studies). There is a growing body of research into the training aspects of end-users, showing the effectiveness, or ineffectiveness, of awareness campaigns, the usability of serious games and so on. The contribution of behavioural science to the field of security by design is not as formalised yet. Therefore, this rapid review has been commissioned by NCSC-NL to gain insight in best practices and potential future research avenues so that behavioural science can be integrated in broader security by design methodologies and projects.

As this rapid review focuses solely on the behavioural aspect of security by design, we refer to this academic field as *security by behavioural design*. The aim of security by behavioural design is to design systems in such a way that the user of these systems is more likely to behave in a secure manner. The goal of this rapid review is to cover the research that empirically tests the effectiveness of various methods, rather than focusing on the more theoretical and often hypothetical research, where solutions are proposed but not tested. In addition to reporting the literature findings, we shared some general insights from the rapid review with some experts in the field to examine their views of the viability of behavioural solutions through a sense check. Lastly, suggestions for future research are provided. In the appendix, we provide an annotated bibliography of a set of theoretical papers that we have found while carrying out this rapid review.

1.2 Choice architecture and nudge theory

Thaler and Sunstein[1], popularised the term *nudging* to refer to the practice of influencing individuals' behaviour by restructuring the choices that are presented to them, a practice also known as *choice architecture*. The goal of nudging is to, often in subtle ways, encourage individuals to make a choice that is preferred by the choice architect while protecting personal freedom of choice. Individuals are not obliged to make a certain decision, nor are alternatives completely removed, but the options they can choose from are selected and presented in such a way that the 'preferred' option is more likely to be chosen. It is therefore important to keep in mind that nudging does usually not lead to a 100% compliance

rate. It does, however, ensure that the way in which choices are offered is the optimal method from the choice architect's point of view, leading to the highest level of compliance without the need for punishment, or restricting individuals in their freedom of choice.

There are various techniques that fall under the umbrella of nudging. In this rapid review, we take a broad view of nudging in general, so that any relevant research in this area can be included. Some categories of nudging include defaulting, opt-in or opt-out systems, providing social norms, and deciding how and in what order to present options. Each of these examples is discussed below.

Defaulting. In defaulting, the preferred option is made the default, while alternative options are still available. The effectiveness of this approach lies in the inertia of the target group, that does not want to put in the effort of making a different choice. Defaulting seems to be easy to implement in software and might be a successful approach to stimulate security behaviours.

Opt-in versus opt-out. This category be viewed as a special case of defaulting in cases where people must decide on a binary outcome. They either perform the behaviour, or they do not perform the behaviour. Opting-in refers to the target group having to act, in order to *agree* with the decision (e.g., clicking to accept cookies), while opting-out refers to the target group needing to act, in order to *disagree* with the decision (e.g., clicking to reject cookies). A well-known real-life example of this is the Dutch government's policy regarding organ donation. This used to be an opt-in system but has recently been turned into an opt-out system. The effectiveness, as with the more generic defaulting approach, lies in the inertia of the target group, which ensures that people who do not act, make the choice that is preferred by the designer of the software.

Social norms. Social norms refer to making salient what most people choose, in order to nudge people to act in a similar fashion. Studies have shown that if people do not know which option to choose, an indication that one option is chosen 'most often' makes them more likely to choose that option themselves. Social norms are generally a powerful tool to nudge individuals towards compliance through the behaviour of the majority. However, using social norms can backfire if the majority is choosing an option that is not to be preferred, as could be the case in a setting where most people click 'accept all' when presented with a cookie statement.

Number of choices and presentation. As a form of a 'rest' category, there are other insights regarding how choices are offered, that are not as distinctly defined as the previous categories. For example, there is the 'decoy' effect, where an unenticing option is added to a list of choices, to make some of the other (preferred) choices stand out more positively. Similar, some researchers found that when lining up options from left to right that are roughly equally enticing, people are slightly more prone to choose the option at the right end of the line, compared to the other options. Furthermore, options can be framed as gains and losses (are we gaining or losing security by performing a behaviour?) which can also influence choices made by end-users. In this rapid review, we include all options discussed here, as well as any other nudge techniques that may be investigated in the scientific literature on security by behavioural design.

Method

2.1 Rapid review

2.1.1 Search strategy and results

We searched the scientific databases PubMed, PsycInfo and Web of Science for relevant publications using the following set of search terms: (cyber* OR *security OR privacy) AND (nudg* OR choice OR option OR opt-in OR opt-out OR norm* OR default* or salien*). Entries were included if the search terms were included in the title of the article. The asterisk is included to ensure that all spellings of the search terms are included. For instance, nudg* ensures that 'nudge', 'nudges' and 'nudging' are all considered relevant by the search engine. Beyond these three scientific databases, we also search Google Scholar and two conference proceedings databases (ACM Digital Library and DBLP Computer Science Bibliography) for relevant entries. The search resulted in 1451 entries, of which 112 were duplicates. After checking the remaining 1339 for relevance based on title, 102 articles remained. A full-text check resulted in a further 56 articles being removed, the majority (45/56) for not covering nudging research. Of the remaining 46 articles, 10 articles were theoretical rather than empirical. In Appendix A, we provide an annotated bibliography of these articles, but do not analyse them further. The final sample therefore consists of 36 articles describing empirical research on security by behavioural design.

2.1.2 Analysis strategy

The 36 papers are categorised based on type of cybersecurity behaviour that is investigated. For each topic, the main findings across the included papers are summarised. This includes which methods seem to be effective in improving cybersecurity behaviour, as well as the nudges that do not seem to be effective in this matter. Furthermore, the setting of the studies is characterised, as there are differences in applicability between studies that have been conducted in lab settings, field settings, or in other settings.

2.2 Sense check

To ensure that the rapid review is not merely a theoretical exercise, we also conducted a brief sense check. For this sense check, experts in the field were contacted and provided with a set of questions regarding the general concepts of security by behavioural design. The questions that were posed, were designed in consultation with the NCSC and aimed at the feasibility and usability of taking nudges into account when designing software (see Appendix B for an overview of the questions). After consultation with the NCSC regarding which experts to include in the sense check, we contacted eight individuals who might be able and willing to take part from academia and the public and private sector. Four responses were collected, and the findings are summarised in the results section. The responses came from employees from academia and the public sector, and respondents worked at organisations that could be considered large (all organisations were larger than 300 FTE).

Results

3.1 Rapid Review

3.1.1 Overview of topics

The included articles cover a range of cybersecurity behaviours. The majority of articles focused on privacy related behaviours. This could refer to privacy settings in various applications, as well as disclosing personal information online, or to what extent people are willing to pay extra for more privacy. A second behaviour that was investigated relatively often is that of passwords. Most of these papers focused on methods to nudge people towards creating stronger passwords for their accounts. The remaining papers focused on other topics such as wireless networks, security choices (e.g. which Wi-Fi network to connect with) and phishing.

3.1.2 Privacy

Privacy was the most investigated topic in the scientific literature on nudging cybersecurity. A range of nudges and settings were used, and privacy was operationalised in various ways. For instance, framing[2–6], defaulting[3,7–9], options provided[10,11], opt-in & opt-out[12–14], salience[15–17], time delays before sharing information[18,19], and social norms nudges[5,20–25] were used to improve privacy behaviour. All these different types of nudges seem to be effective in improving privacy behaviour. However, the studies were carried out in controlled environments, often using panel data such as Amazon's Mechanical Turk (MTurk). Studies incorporated mock-ups of apps or app stores to test effectiveness in these settings.

The specific privacy behaviours that the studies attempted to improve different greatly. Topics included making decisions about which apps to download and install[4,6], in-app privacy settings[2,3,7,12,15] recalling privacy notifications[16], paying for privacy friendly options or apps[8], and disclosing personal information[5,10,17,22]. To further explain the range of research on privacy nudging, we will cover three examples below.

A first example is an online study that investigated the effect of the use of defaults in sharing personal data on the installation of a birthday app[9]. If irrelevant data was 'checked' so that it would be shared with the app maker in case of installing the app, this reduced install rates compared to a 'no-boxes-checked' control condition and a condition in which only relevant data was shared (own and friends' birthdays). This effect was even stronger when all data was 'checked' to be shared upon installation.

A second example is a study that investigated whether adding a privacy indicator in a mock-up app store would influence which apps participants chose to download for specific purposes (e.g. weather predictions, video calling)[6]. While the privacy indicator influenced the likelihood of participants installing an app, with apps being less likely to be installed if the privacy indicator showed low privacy protection, this effect was attenuated when participants were familiar with an app. In those cases, the familiarity was strong enough to overcome a negative privacy indicator.

A third and final example is a study on privacy settings on social media[7]. This study showed that when defaulting privacy settings regarding who could view the participants' content, the default of "only me", or "everyone" both successfully nudged people into lower and higher sharing levels respectively. Interestingly, not only did people choose the default more often, they also were more likely to choose an option that was close to the default if they wished to change the setting in the first place, an effect that can be linked to the

anchoring heuristic literature (the following options were provided: “Everyone,” “Friends of Friends,” “Friends,” “Close Friends,” and “Only Me”).

3.1.3 Passwords

Seven papers investigated the possibility of nudging end-users towards improved password strength[26–32]. One of these studies proposed, but did not test, a nudge to improve password strength, suggesting that mnemonics could be of use when end-users need to come up with answers to security questions[31]. All studies ran online (where reported), and often used platforms such as Prolific and Mturk. All studies typically tested the nudge against a control group, or used a pretest-posttest design, where participants first created a password without being presented with a nudge, then received a nudge and again created a password. The outcome was usually measured using an algorithm to calculate a ‘password strength score’. Generally, nudging seems to be effective, as five studies showed improvements in password strength[26–29,32] although a full experimental design was lacking in some studies. Therefore, in those studies it is unclear whether the password strength increased as a result of the nudge, or for other reasons such as practice throughout the study, or other external factors. One study using the decoy effect did not find a significant improvement in password strength, potentially due to the small sample size and vague definition of the decoy used in that study[30]. The effective studies showed that guidance in creating passwords[26], password meters[26–29,32] or radars[27], and ‘crack time’ estimators[29] are all promising nudges to improve password strength, with studies showing mixed results as to whether personalising a nudge (selecting a specific nudge based on individual characteristics) would be more effective[27–29].

3.1.4 Wireless networks

Three studies investigated the use of nudging in relation to wireless networks [32–34]. One of these studies was a proof of concept using a small number of participants (N =18) which showed the effectiveness of a newly designed wireless network setup wizard in improving security settings[33]. This study showed that using defaults and providing options for naming the network resulted in more secure settings. The other studies focused on redesigning the Wi-Fi selection screen and tested this on university staff and students[34] and on MTurk participants[32]. In MTurk participants, ordering the Wi-Fi networks from most to least secure was successful in nudging participants towards more secure Wi-Fi networks. This was not the case for the university staff and students. In that study, merely ordering the Wi-Fi networks from most to least secure did not result in improved Wi-Fi network choices

in university staff and students. However, using colour coding (traffic light system), and combining colour coding with ordering the Wi-Fi networks did result in more secure Wi-Fi networks being chosen by end-users suggesting that nudging people towards using more secure Wi-Fi connection is feasible.

3.1.5 Other security behaviours

A small number of papers covered other security behaviours: cookie acceptance[35], phishing[36], choice of cloud service[32], and smartphone encryption choices[32]. Coventry and colleagues showed in an online study that a cookie statement that included a minority social norm (where participants were told that 37% of users similar to them accepted the cookies) was effective in reducing the chance of people accepting these cookies themselves[35]. Interviewing participants about the potential effects of a quiz question relating to an email they are about to open found that participants felt this would be effective in preventing phishing attacks[36]. The study on choice of cloud service and smartphone encryption choices found that a social norm ("most popular" banner above preferred option) and defaulting ("yes" to encryption) increased the likelihood of participants choosing for secure cloud services and smartphone encryption respectively[32].

3.2 Sense check

At the start of the sense check, experts were asked to describe what they think 'security by design' means. The experts seem to agree on the definition of security by design and the benefits of implementing security by design principles when they design software. One of the experts defined security by design as *"the practice of acknowledging cyber risk and its mitigation early in the stages of producing IT (both hard- and software). Security by design infers that cyber risk and mitigation is a qualitative aspect during the whole creation cycle of IT systems and services and this includes software design"*, a definition that seems to be echoed by the others. Security by design is gaining ground, and some experts explicitly state that they incorporate security by design principles in their software design.

Respondents state that, when developing software, they consider the potential for end-users to behave in an unsecure manner, but the ways in which they do varies. For example, one expert explained that any end-user risk would come up in a risk analysis of a system that is being designed, and that action would then be taken. Another expert stated the importance of taking into account insider threat, but also user error, both accidental and as the result of a social engineering attack. The behavioural components of the software are sometimes tested

for effectiveness, but a structured, regular, systematic test of these components is often lacking.

The experts are unanimous in that they believe that nudging could be a useful tool in improving cybersecurity. They are aware of the complex nature of human behaviour and therefore believe that it requires a thorough understanding of the users' mindset. As a result, nudging in software design is currently mostly used for 'low hanging fruit' as applying nudges to these behaviours is most likely to work without needing a thorough understanding of users. The experts do not see a specific area where nudging might be useful but point out that nudging could be applied to a wide range of cybersecurity topics, including phishing, configuration choices and data classification systems, but also more generally whenever end-users are required to choose between security and productivity.

While the experts see the use of nudging, there are some concerns about ethics. One expert explained the importance of transparency, pointing out that end-users should not regret choices they made had they been presented in a different manner while another expert pointed out concerns from the public. Furthermore, when asked about downsides of nudging security over productivity, experts point out the need to check the effectivity and efficiency of the nudges included in the software, so that informed decisions can be made about whether to include the nudges in the software design.

In addition to nudging, we also asked the experts about techno-regulation. Techno-regulation is a subfield of law, which suggests that security can be forced by taking away the freedom to act differently. Compared to nudging, this means not merely suggesting a course of action, but preventing end-users from doing anything that is not the preferred option from the software developers' point of view. The major upside is that behaviours that would lead to cyber insecurity are not possible to perform within the software. However, the downside is the potential to work around the software, which reduces the monitoring possibilities.

The experts see potential in using techno-regulation and nudging alongside each other. Techno-regulation can be used for the high-risk behaviours where removing options is justified, whereas nudging can be implemented in cases where the risk is low, or when systems are used for a variety of tasks, each with their own risk level. Furthermore, one experts pointed out the necessity of understanding when an option is 'reliably unwanted' to be available for the end-user. If this information is available, software could be developed in such a way that when the option is reliably unwanted, it is also inaccessible to the end-user, whereas otherwise the option can be accessed.

Discussion

4.1 General findings

The results from this rapid review suggest that nudging is an effective tool to positively influence cybersecurity behaviours. As is generally the case with nudges, they do not seem to be long term solutions for cybersecurity issues, as shown in research on a range of cybersecurity behaviours such as picking a secure Wi-Fi network, or creating a strong password[32] and as pointed out by some of the experts in the sense check. This implies that a nudge should be presented every time an end-user is to make a choice between security and usability. The long-term effects of this repeated nudge presentation are not yet known, and it is possible that the effectiveness diminishes over time. A potential solution for this problem might be to design new nudges throughout the updating process or include variations on the basic nudge that capture the attention of users and reinforce the effect of the nudges implemented in the software. This would especially be of interest when applying social nudges, compared to more static nudges such as defaulting.

The support from experts to include insights from nudge theory in software design is promising. There are some barriers to overcome such as how to ensure nudges are implemented in an ethical manner and ensuring that developers thoroughly test the effectiveness of the proposed nudges to avoid unwanted side-effects. Furthermore, closer collaborations with behavioural scientists might be needed to implement nudges relating to more complex decisions or user behaviours, or to apply effective nudges in complex environments.

4.2 Generalisability of methods

Most articles covered in the rapid review investigated nudging options for privacy and password related behaviours. The question remains to what extent these findings can be generalised to other cybersecurity behaviours as well. For example, in practice, organisations are spending a large part of their resources on phishing training and testing their employees. While some researchers have hypothesised that nudging could also be effective in improving phishing detection[36] and this was also mentioned by experts in the sense check, research in this area is lacking thus far. Nudging could be a promising solution to other forms of social engineering such as tailgating but this has not been investigated yet.

4.3 Areas for future research

Based on the rapid review, it is clear that research on security by behavioural design mostly investigates its effectiveness in a controlled environment. While this did not often take the form of a lab study, the controlled environment was ensured by presenting participants with online surveys, mock-ups of devices or applications, or hypothetical situations in which participants had to make security related decisions. Future research is needed to further test the feasibility of the methods in more realistic settings. On the one hand, this can include more realistic lab studies (e.g. as part of a larger study on workflow or teamwork) so that the environment can be controlled, yet the ecological validity can be protected. On the other hand, this could also entail testing some of the more promising findings in a real-world setting, where the trade-off between security and usability is more visible, and the choices made could have direct consequences on workflow and the security of the organisation.

Based on the sense check, more structured and regular testing of the effectiveness of nudges is needed. The experts stated that they do sometimes run some tests (e.g., an a/b-test to see which user interface is more effective), but that this is not a structural part of the design process. This hinders the creation and collection of best practices, as it remains unclear how much nudging can or does contribute to increasing security. Deeper collaborations with behavioural scientists could be part of the solution. The experts pointed out that in its current form, they would opt for using nudging only for the low hanging fruit, as the human mind is considered complex. Deeper collaborations with behavioural scientists could help to design and test nudges that improve cybersecurity in more complex systems and environments. For example, behavioural scientists could assist in designing nudges that are required to be seamlessly woven into the existing workflow or help counter potential shadow IT behaviours by end-users (or other unwanted behavioural side-effects) in those cases where nudging is not an option, and techno-regulation is applied instead.

One area in which NCSC-NL could support organisations when designing and implementing nudges, is in developing nudging standards. Broadly speaking, these standards could cover three aspects: 1) When to implement nudges, and when to use techno-regulation to ensure end-users behave in a secure way; 2) Guidelines on ethical standards for nudging to improve cybersecurity; and 3) Developing best practices and tools to help organisations choose relevant nudges for the behavioural cyberthreats they are attempting to mitigate.

Nudging has shown to be a promising avenue to improve cybersecurity within organisations when techno-regulation is unwanted or simply not feasible. It cannot fully

replace other types of defences but has shown to be of importance in the cybersecurity field. Including security by behavioural design in the broader security by design context will help organisations to become more secure.

References

1. Thaler RH, Sunstein CR. *Nudge*. Penguin Group; 2009.

2. Adjerid I, Acquisti A, Loewenstein G. Choice Architecture, Framing, and Cascaded Privacy Choices. *Manag Sci*. 2019;65(5):2267-2290. doi:10.1287/mnsc.2018.3028

3. Bahirat P, Sun Q, Knijnenburg BP. Scenario Context v/s Framing and Defaults in Managing Privacy in Household IoT. In: *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion*. IUI '18 Companion. Association for Computing Machinery; 2018. doi:10.1145/3180308.3180372

4. Choe EK, Jung J, Lee B, Fisher K. Nudging People Away from Privacy-Invasive Mobile Apps through Visual Framing. In: Kotzé P, Marsden G, Lindgaard G, Wesson J, Winckler M, eds. *Human-Computer Interaction - INTERACT 2013 - 14th IFIP TC 13 International Conference, Cape Town, South Africa, September 2-6, 2013, Proceedings, Part III*. Vol 8119. Lecture Notes in Computer Science. Springer; 2013:74-91. doi:10.1007/978-3-642-40477-1_5

5. Barev TJ, Schwede M, Janson A. The Dark Side of Privacy Nudging - An Experimental Study in the Context of a Digital Work Environment. In: *54th Hawaii International Conference on System Sciences, HICSS 2021, Kauai, Hawaii, USA, January 5, 2021*. ScholarSpace; 2021:1-10. http://hdl.handle.net/10125/71117

6. Bock S, Momen N. Nudging the User with Privacy Indicator: A Study on the App Selection Behavior of the User. In: *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*. NordiCHI '20. Association for Computing Machinery; 2020. doi:10.1145/3419249.3420111

7. Cho H, Roh S, Park B. Of promoting networking and protecting privacy: Effects of defaults and regulatory focus on social media users' preference settings. *Comput Hum Behav*. 2019;101:1-13.

8. Dogruel L, Joeckel S, Vitak J. The valuation of privacy premium features for smartphone apps: The influence of defaults and expert recommendations. *Comput Hum Behav*. 2017;77:230-239. doi:10.1016/j.chb.2017.08.035

9. Wang N, Wisniewski P, Xu H, Grossklags J. Designing the Default Privacy Settings for Facebook Applications. In: *Proceedings of the Companion Publication of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*. CSCW Companion '14. Association for Computing Machinery; 2014:249-252. doi:10.1145/2556420.2556495

10. Knijnenburg BP, Kobsa A, Jin H. Preference-Based Location Sharing: Are More Privacy Options Really Better? In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '13. Association for Computing Machinery; 2013:2667-2676. doi:10.1145/2470654.2481369

11. Zimmerman S, Thorpe A, Fox C, Kruschwitz U. Privacy Nudging in Search: Investigating Potential Impacts. In: *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval*. CHIIR '19. Association for Computing Machinery; 2019:283-287. doi:10.1145/3295750.3298952

12. Baek YM, Bae Y, Jeong I, Kim E, Rhee JW. Changing the default setting for information privacy protection: What and whose personal information can be better protected? *Soc Sci J*. 2014;51(4):523-533. doi:10.1016/j.soscij.2014.07.002

13. Lai Y-L, Hui K-L. Internet Opt-in and Opt-out: Investigating the Roles of Frames, Defaults and Privacy Concerns. In: *Proceedings of the 2006 ACM SIGMIS CPR Conference on Computer Personnel Research: Forty Four Years of Computer Personnel Research: Achievements, Challenges & the Future*. SIGMIS CPR '06. Association for Computing Machinery; 2006:253-263. doi:10.1145/1125170.1125230

14. Warberg L, Acquisti A, Sicker D. Can Privacy Nudges Be Tailored to Individuals' Decision Making and Personality Traits? In: *Proceedings of the 18th ACM Workshop on Privacy in the Electronic Society*. WPES'19. Association for Computing Machinery; 2019:175-197. doi:10.1145/3338498.3358656

15. Almuhimedi H, Schaub F, Sadeh N, et al. Your Location Has Been Shared 5,398 Times! A Field Study on Mobile App Privacy Nudging. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. Association for Computing Machinery; 2015:787-796. doi:10.1145/2702123.2702210

16. Balebako R, Schaub F, Adjerid I, Acquisti A, Cranor L. The Impact of Timing on the Salience of Smartphone App Privacy Notices. In: *Proceedings of the 5th Annual ACM CCS Workshop on Security and Privacy in Smartphones and Mobile Devices*. SPSM '15. Association for Computing Machinery; 2015:63-74. doi:10.1145/2808117.2808119

17. Hermstruewer Y, Dickert S. Sharing is daring: An experiment on consent, chilling effects and a salient privacy nudge. *Int Rev LAW Econ*. 2017;51:38-49. doi:10.1016/j.irle.2017.06.001

18. Wang Y, Leon PG, Scott K, Chen X, Acquisti A, Cranor LF. Privacy Nudges for Social Media: An Exploratory Facebook Study. In: *Proceedings of the 22nd International Conference on World Wide Web*. WWW '13 Companion. Association for Computing Machinery; 2013:763-770. doi:10.1145/2487788.2488038

19. Wang Y, Leon PG, Acquisti A, Cranor LF, Forget A, Sadeh N. A Field Trial of Privacy Nudges for Facebook. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '14. Association for Computing Machinery; 2014:2367-2376. doi:10.1145/2556288.2557413

20. Masaki H, Shibata K, Hoshino S, Ishihama T, Saito N, Yatani K. Exploring Nudge Designs to Help Adolescent SNS Users Avoid Privacy and Safety Threats. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. Association for Computing Machinery; 2020:1-11. doi:10.1145/3313831.3376666

21. Mendel T, Toch E. Susceptibility to social influence of privacy behaviors: Peer versus authoritative sources. In: *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ; 2017:581-593.

22. Ortloff A-M, Zimmerman S, Elsweiler D, Henze N. The Effect of Nudges and Boosts on Browsing Privacy in a Naturalistic Environment. In: *Proceedings of the 2021 Conference on Human Information Interaction and Retrieval*. CHIIR '21. Association for Computing Machinery; 2021:63-73. doi:10.1145/3406522.3446014

23. Spottswood EL, Hancock JT. Should I Share That? Prompting Social Norms That Influence Privacy Behaviors on a Social Networking Site. *J Comput-Mediat Commun*. 2017;22(2):55-70. doi:10.1111/jcc4.12182

24. Zarouali B, Poels K, Ponnet K, Walrave M. The influence of a descriptive norm label on adolescents' persuasion knowledge and privacy-protective behavior on social networking sites. *Commun Monogr*. 2021;88(1, SI):5-25. doi:10.1080/03637751.2020.1809686

25. Zhang B, Xu H. Privacy Nudges for Mobile Applications: Effects on the Creepiness Emotion and Privacy Attitudes. In: *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. CSCW '16. Association for Computing Machinery; 2016:1676-1690. doi:10.1145/2818048.2820073

26. Furnell S, Alotaibi F, Esmael R. Aligning Security Practice with Policy: Guiding and Nudging towards Better Behavior. In: Bui T, ed. *52nd Hawaii International Conference on System Sciences, HICSS 2019, Grand Wailea, Maui, Hawaii, USA, January 8-11, 2019*. ScholarSpace; 2019:1-10. http://hdl.handle.net/10125/59998

27. Hartwig K, Reuter C. Nudging users towards better security decisions in password creation using whitebox-based multidimensional visualisations. *Behav Inf Technol*. Published online January 22, 2021. doi:10.1080/0144929X.2021.1876167

28. Hupperich T, Dassel K. On the Usefulness of User Nudging and Strength Indication Concerning Unlock Pattern Security. In: Wang G, Ko RKL, Bhuiyan MZA, Pan Y, eds. *19th IEEE International Conference on Trust, Security and Privacy in Computing and Communications, TrustCom 2020, Guangzhou, China, December 29, 2020 - January 1, 2021*. IEEE; 2020:1646-1654. doi:10.1109/TrustCom50675.2020.00227

29. Peer E, Egelman S, Harbach M, Malkin N, Mathur A, Frik A. Nudge me right: Personalizing online security nudges to people's decision-making styles. *Comput Hum Behav*. 2020;109. doi:10.1016/j.chb.2020.106347

30. Seitz T. *Supporting Users in Password Authentication with Persuasive Design*. Tobias Seitz; 2018.

31. Senarath A, Arachchilage NAG, Gupta BB. Security Strength Indicator in Fallback Authentication: Nudging Users for Better Answers in Secret Questions. *CoRR*. 2017;abs/1701.03229. http://arxiv.org/abs/1701.03229

32. Zimmermann V, Renaud K. The Nudge Puzzle: Matching Nudge Interventions to Cybersecurity Decisions. *Acm Trans Comput-Hum Interact*. 2021;28(1):7. doi:10.1145/3429888

33. Ho JT, Dearman D, Truong KN. Improving Users' Security Choices on Home Wireless Networks. In: *Proceedings of the Sixth Symposium on Usable Privacy and Security*. SOUPS '10. Association for Computing Machinery; 2010. doi:10.1145/1837110.1837126

34. Turland J, Coventry LM, Jeske D, Briggs P, Moorsel APA van. Nudging towards security: developing an application for wireless network selection for android phones. In: Lawson SW, Dickinson P, eds. *Proceedings of the 2015 British HCI Conference, Lincoln, United Kingdom, July 13-17, 2015*. ACM; 2015:193-201. doi:10.1145/2783446.2783588

35. Coventry LM, Jeske D, Blythe JM, Turland J, Briggs P. Personality and Social Framing in Privacy Decision-Making: A Study on Cookie Acceptance. *Front Psychol*. 2016;7. doi:10.3389/fpsyg.2016.01341

36. Vitek V, Syed Shah T. *Implementing a Nudge to Prevent Email Phishing*.; 2019.

Appendix A: Annotated bibliography of theoretical papers

(Acquisti et al., 2017)
Acquisti and colleagues provide an overview of research into privacy and decision-making based on literature that relates to nudging theory and choice architecture, discussing the main findings and challenges.

(Ariely & Holzwarth, 2017)
Ariely and Holzwarth discuss choice architecture in relation to decision-making processes in the privacy domain, focusing on trade-offs.

(Coventry et al., 2014)
Coventry and colleagues offer a structured approach to implementing nudges in cybersecurity settings. They build on the MINDSPACE approach and develop their approach to be of specific use in the cybersecurity domain.

(Dogruel, 2019)
Dogruel conducted interviews with German and US citizens to discuss their preferences for System 1 (heuristic decision-making) and System 2 (deliberate decision-making) type nudges, suggesting that they preferred System 2 over System 1 style nudges.

(Figl & Lehrer, 2020)
Figl and Lehrer presented a proposed study into the influence of various privacy setting options on disclosure behaviour by end-users. The findings of this study are not yet available.

(Noain Sanchez, 2016)
Noain-Sánchez argues for the need of privacy by default and how opt-in systems could be applied to privacy settings in online environments such as social network sites.

(Renaud & Zimmermann, 2018)
Renaud and Zimmerman discuss the ethics of nudging and offer guidelines that can help decide whether a nudge is ethically justifiable. The paper specifically focuses on ethical factors that relate to the themes of information security and privacy.

(Toch et al., 2010)
Toch and colleagues presented a work-in-progress at CHI 2010 in which they built a system to present users with specific privacy settings based on a range of parameters, suggesting that tailored settings could be presented to users, instead of a one-size-fits-all.

(Westin & Chiasson, 2019)
Westin and Chiasson discuss dark patters, nudging techniques used to make users more likely to make choices that go against their best interest online (e.g. cookie walls where accepting all cookies is easier than rejecting all cookies). They suggest that a 'FoMO-Centric Design' can explain the effectiveness of these dark patterns.

(Ziegeldorf et al., 2015)
In this position paper Ziegeldorf and colleagues propose a design paradigm for privacy that considers the human tendency to compare. In this paradigm users can compare their settings to those of specifically selected comparison groups and decide whether they would want stricter or less strict privacy options in comparison with these target groups. They call this paradigm 'Comparison-based Privacy' (CbP).

References annotated bibliography

Acquisti, A., Adjerid, I., Balebako, R., Brandimarte, L., Cranor, L. F., Komanduri, S., Giovanni Leon, P., Sadeh, N., Schaub, F., Sleeper, M., Wang, Y., & Wilson, S. (2017). Nudges for Privacy and Security: Understanding and Assisting Users' Choices Online. *Acm Computing Surveys*, *50*(3), 44. https://doi.org/10.1145/3054926

Ariely, D., & Holzwarth, A. (2017). The choice architecture of privacy decision-making. *Health and Technology*, *7*(4), 415–422. https://doi.org/10.1007/s12553-017-0193-3

Coventry, L. M., Briggs, P., Jeske, D., & Moorsel, A. P. A. van. (2014). SCENE: A Structured Means for Creating and Evaluating Behavioral Nudges in a Cyber Security Environment. In A. Marcus (Ed.), *Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience—Third International Conference, DUXU 2014, Held as Part of HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014, Proceedings, Part I* (Vol. 8517, pp. 229–239). Springer. https://doi.org/10.1007/978-3-319-07668-3_23

Dogruel, L. (2019). Privacy nudges as policy interventions: Comparing US and German media users' evaluation of information privacy nudges. *Information, Communication & Society*, *22*(8), 1080–1095.

Figl, K., & Lehrer, C. (2020). Privacy Nudging: How the Design of Privacy Settings Affects Disclosure in Social Networks. In F. Rowe, R. E. Amrani, M. Limayem, S. Newell, N. Pouloudi, E. van Heck, & A. E. Quammah (Eds.), *28th European Conference on Information Systems—Liberty, Equality, and Fraternity in a Digitizing World, ECIS 2020, Marrakech, Morocco, June 15-17, 2020*. https://aisel.aisnet.org/ecis2020_rip/79

Noain Sanchez, A. (2016). Knowledge as an effective tool to protect ICT users' privacy. The layered informed consent as `opt-in' model. *Doxa Communicacion*, *22*, 149–155.

Renaud, K., & Zimmermann, V. (2018). Ethical guidelines for nudging in information security & privacy. *International Journal of Human-Computer Studies*, *120*, 22–35. psyh. https://doi.org/10.1016/j.ijhcs.2018.05.011

Toch, E., Sadeh, N. M., & Hong, J. (2010). Generating Default Privacy Policies for Online Social Networks. *CHI '10 Extended Abstracts on Human Factors in Computing Systems*, 4243–4248. https://doi.org/10.1145/1753846.1754133

Westin, F., & Chiasson, S. (2019). Opt out of Privacy or 'Go Home': Understanding Reluctant Privacy Behaviours through the FoMO-Centric Design Paradigm. *Proceedings of the New Security Paradigms Workshop*, 57–67. https://doi.org/10.1145/3368860.3368865

Ziegeldorf, J. H., Henze, M., Hummen, R., & Wehrle, K. (2015). Comparison-Based Privacy: Nudging Privacy in Social Media (Position Paper). In J. García-Alfaro, G. Navarro-Arribas, A. Aldini, F. Martinelli, & N. Suri (Eds.), *Data Privacy Management, and Security Assurance—10th International Workshop, DPM 2015, and 4th International Workshop, QASA 2015, Vienna, Austria, September 21-22, 2015. Revised Selected Papers* (Vol. 9481, pp. 226–234). Springer. https://doi.org/10.1007/978-3-319-29883-2_15

Appendix B: Sense check questions

We are currently conducting a research project into security by behavioural design, where we look at the scientific literature around 'nudging' in relation to cybersecurity. Nudging (also known as choice architecture) is a method of steering end-users' behaviour by means of presenting choices or options in a specific way.

Some examples of nudging are:

1. Defaulting, where the preferred option is the default (e.g. a ticked box in a form)
2. Social proof, where users are informed of the behaviour/choices of peers (e.g. informing users of which settings their peers usually accept)
3. Direct feedback, where the consequences of various choices are presented (e.g. a 'password meter' that shows the strength of a newly generated password

To complement the findings from the literature review, we are performing a 'sense check' where we ask experts in the field about their opinions and experiences. Below, we outline 7 questions that we would like to get you to answer in either English or Dutch. The answers can be as short or long as you like, and you can respond by simply replying to this email. If any questions are unclear, please feel free to contact us.

1. How would you define "security by design" and is it something that you apply in software development processes?

2. When developing software, do you take into account the potential risky (unsecure) behaviour of end-users, and if so: how?

3. Do you measure the effectiveness of your behavioural components (for example, through a/b-testing)?

4. We would like to ask you whether you think nudging techniques to influence end-users to act more securely could be useful.

5. Nudging has been applied to various cybersecurity behaviours, such as creating stronger passwords or updating your privacy settings. Do you see other areas of cybersecurity behaviour where these techniques could be useful?

6. One of the downsides of nudging could be that processes are taking longer, or other processes need to be put in place (for instance, the more people are unsure about a phishing email, the more people might contact the information desk which then needs to upgrade its capacity). To what extent do you think these potential downsides might preclude you from using these techniques in your software development?

7. An alternative to nudging is techno regulation, where end-users are not 'nudged' into a specific direction, but unwanted options are removed all together. Would this be a sensible approach for the type of software you develop, if so, how would you see the implementation?

Thank you for your response!