Usable Specification of Security and Privacy Demands

Matching User Types to Specification Paradigms

Manuel Rudolph
manuel.rudolph@iese.fraunhofer.de
Fraunhofer IESE
Kaiserslautern, Germany

Svenja Polst
svenja.polst@iese.fraunhofer.de
Fraunhofer IESE
Kaiserslautern, Germany

Denis Feth
denis.feth@iese.fraunhofer.de
Fraunhofer IESE
Kaiserslautern, Germany

ABSTRACT

Security and privacy are considered important by most users. However, formulating their own abstract data protection requirements is already a challenge for them. The mapping of these requirements to concrete setting options in an application is even more challenging—partially because the user interfaces for data protection settings are not tailored to the needs of different user types. This is one of the reasons why only few users make data protection settings regularly and purposefully. In this paper, we describe different specification paradigms for privacy settings and evaluate which paradigm best suits different user types. We investigate with which paradigm a certain user type achieves the best results in terms of objective and perceived correctness, efficiency and satisfaction.

CCS CONCEPTS

• **Security and privacy** → **Usability in security and privacy**; • **Human-centered computing** → *Interaction design theory, concepts and paradigms.*

KEYWORDS

Privacy Policy Specification, User Types, Specification Interfaces, Correctness, Efficiency, Satisfaction

1 INTRODUCTION

In the Internet, more and more personal data are collected, stored, analyzed, reused and partially sold. For Internet users, it is becoming increasingly complicated to understand and control how their data is used. A majority of users in Germany even feel they have lost control over their personal data and are uncomfortable with this situation [4].

MuC'19 Workshops, Hamburg, Deutschland

https://doi.org/10.18420/muc2019-ws-302-05

To regain control, users need to gain more self-determination. Self-expressing their security and privacy demands for personal data shared with online services is the first step in that direction. Different tools (e.g., privacy dashboards) can support users in expressing their demands. Throughout this paper, we refer to these tools as "Policy Administration Points" (PAPs) and to the expressed privacy demands as "Privacy Policies". In a previous study, we revealed that users currently use PAPs relatively rarely [16]. The study showed that PAPs are being received as too complicated and too time-consuming and users do not feel competent enough for using them. Those and other reasons indicate usability problems with existing PAPs.

In general, PAPs can support different specification paradigms (or paradigms for short). We define a specification paradigm as a pattern that defines the specification process and the user interface in a PAP for the task of policy specification including the expressiveness and the guidance that the user receives during the specification. We assume that different user types experience the usability of a paradigm differently.

In this paper, we present a pre-test for an empirically founded guidance for selecting the appropriate paradigms for a certain user type in terms of sub-qualities of usability, namely (objective and subjective) effectiveness, efficiency and user satisfaction. We used the persona model proposed by Dupree [3] for defining user types and selected four representative paradigms. We conducted an experiment in which we compared the usability improvements when providing best matching paradigms in a PAP to different user types. Our experimental results allowed us to give recommendations for paradigms to be used by a specific user who fits to a given user type.

We already published parts of the raw data to argue on the efficiency and satisfaction [17], as well as on the effectiveness [18] of different paradigms. This paper goes beyond these publications and presents the following new aspects: On the one hand, we extended the statistical analysis of the raw experiment results with a special focus on the user groups. On the other hand, we used the analysis results to recommend concrete paradigms for different user types with respect to usability.

The remainder of this paper is structured as follows: We highlight relevant related work in Section 2. In Section 3, we explain our research questions on matching paradigms to user types for increased usability. We present our empirical work in Section 4 and discuss our findings in Section 5. We conclude in Section 6.

2 RELATED WORK

Privacy Specification Interfaces

In practice, many tools and online services offer privacy settings. These tools implement one or more specification paradigms. Therefore we explored the state of the practice and art in order to derive relevant paradigms.

For example, most modern browsers allow users to enable or disable predefined privacy and security policies. In social media services, users can typically specify their privacy policies in a fine-grained manner on the respective privacy setting pages. Especially large services (e.g., Facebook, Google) often provide specification support (e.g., explanations, examples, template based specification, small wizards) for the specification of security and privacy settings. However, studies reveal that users still misinterpret some of the privacy policies they can specify (e.g., [11]).

In the scientific literature, one can find different approaches and studies regarding PAP usability. The KAoS Policy Administration Tool (KPAT) [21] of the KAoS policy and domain service framework uses a template-based approach with natural English hypertext templates. These templates are specified in an ontology and the tool can transform specified policies into machine-understandable equivalents.

Johnson et al. [8] propose a method and a tool named SPARCLE for eliciting concrete security policies of users with varying background knowledge. The user can enter his security demands in natural language or in a structured natural language-based format. SPARCLE transforms the input into machine-readable policies.

PERMIS [7] is a role-based access control authorization infrastructure allowing users to create policies, for example, via a "Policy Wizard". Thus, this tool provides an interface with sequential, small specification steps in which supportive questions guide the user through the specification process.

Fang and LeFevre [5] present an active learning wizard that supports users in specifying their own privacy and security policies via few brief decisions on whether to share particular information with an entity.

Cranor proposes P3P (Platform for Privacy Preference Project) as a protocol to declare the intended use of information of users on websites in a human-understandable format [1].

User Type and Persona Models

User type and persona models aim for clustering users into categories that explain their character traits and behavior. There are generic models that are not bound to a particular situation or domain, such as the Big Five personality traits [2], Keirsey's Temperaments [9] and the Myers-Briggs Type Indicators [14]. In addition, other approaches relate more specifically to the character traits relevant for security and privacy decisions. Westin conducted more than 30 privacy surveys [10] from which he derived a classification based on users' privacy concerns containing the three categories: Fundamentalist (high concern), Pragmatist (medium concern), and Unconcerned (low concern). However, Urban and Hoofnagel [20] argue Westin's work neglects the role of users' knowledge and available information about privacy practices and domain specific business processes. Smith calculates the privacy concern of a person as a numerical value in his quite generic approach "Concern for Information Privacy (CFIP)" [19]. Malhotra et al. [12] extend previous work (e.g., CFIP) in their approach called Internet Users' Information Privacy Concerns (IUIPC) by reflecting the privacy concerns of Internet users with a special focus on the individuals' fairness perception regarding data privacy. In the approach Information Seeking Preferences [13], Morton clusters users into five groups based on the ranking of 40 privacy related statements. These groups are: Information controllers, security concerned, benefit seekers, crowd followers and organizational assurance seekers. Dupree proposes a privacy persona model [3] that differentiates the users on the basis of two attributes: the user's knowledge about security and privacy on the one hand, and the user's motivation to spend effort to protect privacy and security on the other.

3 MATCHING USER TYPES TO SPECIFICATION PARADIGMS

Our goal is to find the match between a user type and the paradigm that leads to best usability. In this section, we state our research questions, describe characteristics of paradigms and user types and explain the selection criteria for both.

Research Questions

Our fundamental question is: What has to be done so that users successfully use PAPs? In order to refine and answer this questions, we have to consider different aspects. In particular, there can be discrepancies between the resources that are required by a paradigm and resources a certain user type has. This aspect is described in Table 1. These discrepancies can lead to poor usability.

According to ISO 9241 [6], "usability" can be split into the sub-qualities effectiveness, efficiency and satisfaction.

If we want to improve the specification of security and privacy policies, we need to consider two different types of effectiveness: The objectively correct specification of policies with PAPs (objective effectiveness) and the confidence the user has that the specified policy is correct (perceived effectiveness).

Based on this, we phrased the following research questions:

RQ1: Do paradigms significantly differ regarding objective effectiveness for a given user type?
RQ2: Do paradigms significantly differ regarding perceived effectiveness for a given user type?
RQ3: Do paradigms significantly differ regarding efficiency for a given user type?
RQ4: Do paradigms significantly differ regarding user satisfaction for a given user type?

Specification Paradigms

We analyzed existing specification approaches from literature and from PAPs in practical use (cf. Section 2). Based on this analysis, we derived eight specification paradigms that are implemented by the PAPs and approaches. Afterwards, we categorized the paradigms according to their expressiveness (i.e., how many decisions they request) and their guidance (i.e., how much help the user receives during the specification). Based on that, we then selected one representative for all four combinations of high and low expressiveness and guidance. Thus, we came up with the following four paradigms we considered for our evaluation:

Security levels (low expressiveness, high guidance): In this paradigm, the user chooses one out of a limited number of security levels for the entire system. A level contains a set of immutable policies. A "higher" level typically implies higher security (but potentially lower usability).

Default Policies (low expressiveness, low guidance): In this paradigm, for each use case of the system, the user chooses one set of security policies. In contrast to the security level paradigm, there is not necessarily an order implied and the user has to select multiple sets.

Wizard (high expressiveness, high guidance): The user instantiates security and privacy policies in a predefined order within a template-based interface, which is split into several small specification steps. The specification process is well guided in each step.

Template Instantiation (high expressiveness, low guidance): The user instantiates security and privacy policies in a template-based interface in a fine-grained manner. Users can choose the order in which they want to instantiate templates.

For each paradigm, we (subjectively) estimated the resources (cf. Table 1) that are required from a user to apply the paradigm. The results are shown in Table 2. In the future, the required resources should be estimated based on quantitative data.

Table 1: Resources to be considered

Resource	Description
Domain Knowledge	Required and actual knowledge regarding the application domain including the service's use cases for which a policy is to be specified. This knowledge includes information about the personal data that has to be shared with the service. The user needs to understand the domain in order to be capable of making privacy-related decisions.
Security & privacy knowledge	Required and actual knowledge of potential and actual use of personal data by the service and potential threats that arise from this use are necessary in order to be capable of making security and privacy-related decisions. This knowledge also includes that users understand the effect of countermeasures for improving their own security and privacy.
Technical knowledge	Required and actual knowledge of the functionality of the service and its PAP
Time	Required and available time to specify policies in the PAP
Cognitive capacity	Amount of security and privacy related information the user needs vs. is capable of processing simultaneously during the specification of policies in a PAP
Physical capacity	Required and actual accessibility of a device that allows the use of the PAP in the respective system.

Table 2: Requirements on user for different Specification Paradigms

Resource	Defaults	Levels	Template	Wizard
Domain Knowledge	medium	low	high	medium
Security & privacy knowledge	medium	low	high	medium
Technical knowledge	medium	low	high	medium
Time	medium	low	high	medium
Cognitive capacity	medium	medium	high	high

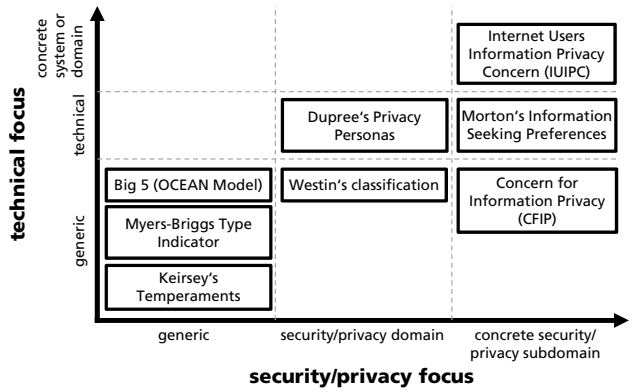

Figure 1: User Type Models

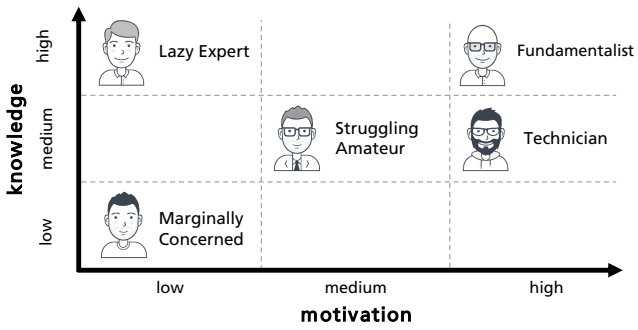

Figure 2: Dupree's Personas (User Types)

User Types

Because users differ in their available resources, we would need to match paradigms to individual users. As this is impractical, we decided to group users according to their resources. We analyzed and characterized existing user type models by the two properties "focus on IT security and privacy" and "focus on technical systems".

There are many models for user types and personas (see Figure 1). They differ in their focus on IT security (y-axis) and privacy as well as in their focus on technical systems (x-axis). In both cases, there are very special models developed for a specific sub-domain or system, but also generic approaches. Since we focus on the specification of privacy settings for technical systems, we have decided against too generic or too special models.

We chose the model by Dupree et al. [3], because it focuses on security and privacy for technical systems, but is not restricted to a too specific sub-domain of security and privacy or to a specific system. This model categorizes users by their motivation and their knowledge to specify security and privacy policies. Dupree has derived the following five personas from personal interviews with 32 university related digital natives. In the following, we consider these personas as representatives of a corresponding user type:

(1) *Marginally Concerned*
(2) *Struggling Amateur*
(3) *Technician*
(4) *Lazy Expert*
(5) *Fundamentalist*

Figure 2 shows the characterization of each user type with respect to knowledge and motivation.

4 EXPERIMENTAL EVALUATION

We performed an experiment for answering our research questions.

Design and Execution

Scenario and Tasks. We derived the scenario and the corresponding privacy policies in our experiment from a project in the context of the digitization of rural areas. In this project, citizens in rural areas have access to digital services such as an online marketplace with local merchants (called BestellBar), a delivery service where citizens deliver goods from local merchants to other citizens (called LieferBar) and a digital village bulletin board (called DorfFunk). The participants were asked to configure privacy policies for these services according to predefined privacy demands, which we provided to every participant in a handout. The presetting of the privacy demands was necessary so that all participants could use the specification interfaces in a comparable way and so that we were able to compare the specified policies to a sample solution. The privacy demands were described in the six task-related statements, one being for example: "When I place an order in the BestellBar app, I do not—under any circumstances—want to receive advertising from other providers that refers to the ordered product. They may not use my data." These demands did not match word by word with the content in the specification interfaces, because a one-to-one match would lead participants to compare buzzwords without thinking about the semantics. The tasks and the short scenario description on the handout were supported by a short video introducing the novel, digital services for citizens of a village.

We created four PAPs, each implementing one of the four paradigms presented before. The participants were asked to specify policies for the same six task-related statements for each paradigm. To provide the content of the PAPs, we derived example privacy policies and corresponding privacy policy templates in a workshop [15] with the developers of the mentioned digital services. The paradigms "template instantiation" and "wizard" allowed participants to instantiate concrete privacy policies from the derived policy templates. The paradigm "default policies" let participants select from a limited list of predefined privacy policies. The paradigm

"security levels" let participants chose one of three different sets of privacy policies. All tasks in the experiment were solvable with all four paradigm implementations.

Procedures and Instruments. Our experiment was implemented as a publicly available online experiment in German and English. Participants could start the experiment once with a unique eight digit participant id. Users were given the opportunity to interrupt the experiment, but not to repeat already finished steps.

Our experiment was structured as follows. First, the participants had to agree to an informed consent and to answer demographic questions about, for example, their age, gender and educational level. Then, a self-assessment followed about one's own expertise and motivation in the areas of security and privacy as well as experience in dealing with digital services. We assessed an potential impact of the participants' characteristics and capabilities on the results of the experiment. Next, the participants had to select one of Dupree's personas [3], which were presented in random order based on nine to twelve character traits formulated in the ego-perspective. The persona names were not displayed in order to reduce the threat of social desirability. Thereupon, the scenario and the concrete tasks were explained via handout and video. In the following steps, the participants were asked to specify all privacy policies according to the given tasks for each of the four paradigms: default policies, security levels, template instantiation and wizard. The paradigms were presented in random order to minimize learning effects. The participants had to rate the current paradigm after each specification regarding perceived effectiveness and user satisfaction. After completing the four specifications, the participants were asked to rank the four specification types according to their preference of using them in real life. Finally, participants should determine how well they were able to identify with the chosen user type and the overall scenario.

Data Analysis. We investigated the research questions as follows.

RQ1: The paradigms offer different levels of expressiveness. Thus, the paradigms required a different number of decisions being made by the participants: one decision for the paradigm "security levels", six decisions for the "default policies", 18 decisions for the "template instantiation" and 18 decisions for the "wizard". Thus, we measured the objective effectiveness as the ratio of incorrect decisions to all decisions.

RQ2: We measured the perceived effectiveness as the self-evaluation with respect to the objective effectiveness. Therefore, we asked the participants after the use of each paradigm whether they think that they solved all tasks correctly (zero mistakes). Then, we calculated the ratio of correct self-evaluation.

RQ3: We measured the elapsed times to perform the policy specifications with the paradigms. We calculated averages of the times per paradigm and per user type.

RQ4: We asked the participants after each specification to rate their satisfaction on a scale from 1 ("I really dislike this specification paradigm") to 5 ("I really like this specification paradigm"). After completion of all four specification rounds, we asked the participants to rank the four specification paradigms according to their preference. We calculated mean and median values per paradigm and per user type.

For measuring the statistical significance, we performed the following tests:

- For RQ1, RQ3 and RQ4, we performed Kruskal-Wallis tests (α= 0.05) to investigate whether the selection of the paradigm has a significant influence on the respective qualities for each user type as well as whether the user type has a significant influence on the quality. We also calculated the effect sizes using Cohen's d value (d_c: small effect: $|d_c|$=0.2; middle effect $|d_c|$=0.5; large effect $|d_c|$=0.8).
- For RQ2, we performed Fisher's exact tests (α= 0.05), which is suitable for small sample sizes, to determine whether the paradigm selection has a significant influence on the perceived effectiveness for each user type as well as whether the user type has a significant influence on the perceived effectiveness. We also calculated the effect sizes using Cramer's φ value (φ_c: for df=3; small effect: $|\varphi_c|$=0.06; middle effect $|\varphi_c|$=0.17; large effect $|\varphi_c|$=0.29).

We had to exclude the user type fundamentalist from some analyzes due to the small number of such participants.

Execution. We invited the participants in the circle of friends and acquaintances of the authors as well as in the authors' institution. We sent 120 personal invitation emails to interested persons with the handout attached, which contained instructions for executing the experiment including the individual participant id. After email sending, we deleted any relation between participant ids and participants in order to ensure anonymity. The online experiment was made available for 14 days and the participation took about 30-40 minutes (no time limit).

Results

In this chapter, the results of our experiment are presented. Table 4 shows the raw results of our experiment.

Participant Description. 61 persons finished the experiment with valid data sets. Their age ranged from 18 to 82 (M=40.54;

SD=14.37). 43 percent of the participants were female. 54 percent of the participants held a university degree, 15 percent a doctoral degree, 11 percent had an entrance qualification for higher education and 18 percent had a secondary school leaving certificate as highest level of education. 54 percent of the participants were related to the authors' institution. Table 3 shows the distribution of user types of the participants.

Objective Correctness. All user types had differences regarding their number of mistakes (objective correctness), when comparing the best to the worst matching paradigm (marginally concerned by 55%, amateurs by 58%, the others by 100%). However, the difference is not significant for every user type. It is significant for the "amateurs" ($H=16.15$, $p<0.01$, $d_c=0.89$), for the "lazy experts" ($H=16.63$, $p<0.01$, $d_c=1.44$) and for the "technicians" ($H=11.15$, $p=0.01$, $d_c=0.86$), but not for the "marginally concerned" ($H=4.98$, $p=0.17$, $d_c=0.43$). Due to the small sample size, the test could not provide significant results for the "fundamentalists". We revealed a significant influence with a large effect of the user type selection on the mistakes made ($H=35.23$, $p<0.01$, $d_c=0.81$). We explain this effect of the user type with the significant difference regarding objective correctness of the marginally concerned compared to the other user types, as they perform significantly worse. We see an influence of the user type selection in each paradigm: "default policies" ($H=13.88$, $p<0.01$), "template instantiation" ($H=14.10$, $p<0.01$), and "wizard" ($H=17.04$, $p<0.01$), and also for the "security levels" ($H=7.99$, $p<0.05$).

Perceived Correctness. All user types had differences regarding the precision in self-evaluation (perceived objectiveness), when comparing the best to with the worst matching paradigm (marginally concerned by 804%, amateurs by 143%, lazy experts by 200%, technicians by 71%; the percentage increase for fundamentalists is infinite). However, the difference is not significant for every user type. It is significant for the "marginally concerned" ($T=12.49$, $p=0.01$, $\varphi_c=0.53$), the "amateurs" ($T=13.78$, $p<0.01$, $\varphi_c=0.41$), and for the "lazy experts" ($T=10.86$, $p=0.01$, $\varphi_c=0.51$), but neither for the "technicians" ($T=4.44$, $p=0.26$, $\varphi_c=0.28$), nor for the "fundamentalists" ($T=6.00$, $p=0.24$, $\varphi_c=0.75$). We found that the selection of the user type also has an influence on the correct self-evaluation ($T=10.08$, $p=0.04$, $\varphi_c=0.20$), but not a very strong one and with only a medium effect size.

Table 3: Chosen user types

User type	Number	Ratio
Marginally Concerned	12	20%
Amateur	21	34%
Lazy Expert	11	18%
Technician	14	23%
Fundamentalist	3	5%
Total	61	100%

Efficiency. All user types had differences regarding the needed time (efficiency), when comparing the most efficient to the least efficient paradigm matching (marginally concerned by 40%, amateurs by 58%, lazy experts by 70%, technicians by 49% and fundamentalists by 69%). However, the difference is not significant for every user type. It is significant for the "amateurs" ($H=23.64$, $p<0.01$, $d_c=1.19$), for the "lazy experts" ($H=13.09$, $p<0.01$, $d_c=1.16$) and for the "technicians" ($H=9.85$, $p=0.02$, $d_c=0.79$), but not for the "marginally concerned" ($H=2.57$, $p=0.46$, $d_c=0.20$). Due to the small sample size, the test could not be meaningfully applied to the "fundamentalists". We did not find a significant effect of the user type selection on the time needed ($H=3.90$, $p=0.27$, $d_c=0.13$). Thus, the distribution of time needed is similar across all user types.

User Satisfaction. All user types had differences regarding the satisfaction, when comparing the best to with the worst matching paradigm (marginally concerned: mean by 1, median by 1; amateurs: mean by 1.7, median by 2; lazy experts: mean by 2.1, median by 2; technicians: mean by 0.9, median by 1; and fundamentalists: mean by 1, median by 1). However, the difference is not significant for every user type. It is significant for the "amateurs" ($H=24.23$, $p<0.01$, $d_c=1.20$), for the "lazy experts" ($H=16.50$, $p<0.01$, $d_c=1.43$), but not for the "marginally concerned" ($H=6.41$, $p=0.09$, $d_c=0.58$) or the "technicians" ($H=4.06$, $p=0.26$, $d_c=0.29$). Due to the small sample size, the test did not yield meaningful results for the "fundamentalists". There was no significant influence of the user type selection on satisfaction ($H=5.87$, $p=0.12$, $d_c=0.23$).

5 DISCUSSION

Matching User Types to Specification Paradigms

The results of our experiment partially answer our four research questions. The marginally concerned performed best (objective effectiveness, perceived effectiveness and efficiency) with the "security levels", but they were unsatisfied with this paradigm. Also, all other personas achieved best objective effectiveness with "security levels". This seems reasonable, as this paradigm requires the least user resources.

The amateurs performed best with respect to perceived effectiveness with "default policies". However, they had comparable results regarding objective effectiveness and efficiency with the paradigms "default policies", "template instantiation" and "wizard".

The lazy experts performed worse than amateurs and technicians in many direct comparisons. This indicates that the motivation of users has a significant influence on the results.

	Defaults	Levels	Template	Wizard	Defaults	Levels	Template	Wizard
	Objective Effectiveness: Ratio of incorrect decisions				Perceived Effectiveness: Ratio of correct self-evaluation			
Marginally Concerned	0.56	0.25	0.49	0.50	0.25	0.75	0.25	0.08
Amateur	0.12	0.05	0.12	0.12	0.62	0.81	0.33	0.33
Lazy Expert	0.15	0.00	0.16	0.21	0.82	0.73	0.27	0.27
Technician	0.17	0.00	0.15	0.11	0.64	0.86	0.57	0.50
Fundamentalist	0.00	0.00	0.06	0.13	1.00	0.67	0.33	0.00
All participants	0.22	0.07	0.20	0.21	0.61	0.79	0.36	0.30
	Efficiency: Average needed time in minutes				Satisfaction: Mean and median of user feedback on scale from 1 (low) to 5 (high)			
Marginally Concerned	4.3	2.6	3.4	4.0	3.3 / 3.5	3.0 / 3	3.9 / 4	4.0 / 4
Amateur	3.4	1.6	3.0	3.8	3.3 / 4	2.1 / 2	3.8 / 4	3.8 / 4
Lazy Expert	2.7	1.1	2.7	3.7	3.0 / 3	1.9 / 2	4.0 / 4	3.8 / 4
Technician	3.5	1.8	3.5	3.5	3.4 / 4	3.2 / 3	4.1 / 4	3.8 / 4
Fundamentalist	3.5	1.4	3.5	4.5	4.3 / 4	3.3 / 4	4.3 / 5	4.3 / 4
All participants	3.5	1.8	3.1	3.8	3.3 / 4	2.6 / 2	4.0 / 4	3.9 / 4

Table 4: Experiment Results

The technicians also achieved best results with "security levels" and "default policies". However, they reached the best results of all personas with respect to objective effectiveness and perceived effectiveness with the paradigms "template instantiation" and "wizard".

We do not draw conclusions about the fundamentalists due to the small number of participants choosing this persona.

In summary, the experiment showed that the matching of paradigms to user types can increase the objective (RQ1) and perceived (RQ2) effectiveness, the efficiency (RQ3), but not significantly for all user types with respect to all qualities. The user satisfaction (RQ4) showed contradictory results, as many participants did not like the paradigm with which they performed best. However, we partially recognized a significant influence of the paradigms on the user types.

Threats to Validity

Internal Validity. We did not control the participants during or after the experiment, but instructed them before and during the experiment as we would have done in a controlled setting. However, we cannot rule out that participants talked about the experiment with other participants before their participation or that participants were distracted during execution. A participant who could not identify with the provided tasks or specified policies well, may have had lower motivation to take effort in correctly using the paradigms, which may negatively affect objective effectiveness. The task of specifying with the paradigm "security levels" in our experiment does most likely not reflect the reality since the preset tasks matched perfectly to one of the security levels. However, we decided to propose a perfect match, because the lack of a perfect solution may have been irritating. Another threat to validity is that the participants selected the user type themselves. Their selection might be prone to social desirability, since certain character traits and attitudes are perceived as desirable. However, the user type descriptions also mention actual behavior, which we assume is less prone to social desirability. We assume that the mixture of attitude and actual behavior in the user type description reduces the risk of social desirability. Also, the persona names (e.g., lazy experts) were not shown in order to reduce this threat.

External Validity. We tried to represent the use of privacy policies in real life. In reality, participants would have their own individual privacy demands. However, we had to provide specific tasks in order to measure the correctness as the discrepancy between the participants' results and the sample solution. Thus, we do not know whether the same results would be achieved in the real world with the participants' own privacy policies. We only base our recommendations on a single experiment and the number of participants per user type was rather small. In addition, a large number of participants were academics, which does not reflect the overall population. Further experiments are necessary for the generalization of our results.

Conclusion Validity. The selection of the paradigms was based on common use in practice and the diverging expressiveness and guidance of the four paradigms. We do not know whether other paradigms would lead to better results in a comparable experiment. This limits the power of our recommendations of best suitable paradigms. Moreover, we cannot distinguish between the findings on paradigms and the user interfaces of the corresponding PAPs. To minimize the influence of the user interface design, we asked usability experts for support in making these tools as unobtrusive as possible.

6 CONCLUSION

In view of the increasing amount of personal-related data that is processed by modern systems, it becomes more and more important for users to express their security and privacy demands and make corresponding settings. As users differ a lot (especially with respect to knowledge, cognitive capacity, available time and motivation), it is important to provide different, complementary PAPs so that a broad range of users can be reached.

In this article, we derived key characteristics of users and paradigms in order to find the best matching of users to paradigms to improve objective effectiveness, perceived effectiveness, efficiency and user satisfaction. Our pre-test provided first evidence for such a match and showed that there are indeed differences between the user types. This means that different paradigms provide different usability for certain user types. Thus, in our future work we aim to explore this match in more detail by extending our experiments. Especially, larger numbers of participants and paradigms will help us to derive a complete picture.

ACKNOWLEDGMENTS

The research presented in this paper has been supported by the German Ministry of Education and Research projects "Nationales Referenzprojekt für IT-Sicherheit in der Industrie 4.0 (IUNO)" (grant no. 16KIS0328) and "Transparente und selbstbestimmte Ausgestaltung der Datennutzung im Unternehmen (TrUSD)" (grant no. 16KIS0898). The sole responsibility for the content of this paper lies with the authors.

REFERENCES

[1] Lorrie Faith Cranor. 2003. P3P: making privacy policies more useful. *IEEE Security Privacy* 1, 6 (2003), 50–55.

[2] J. M. Digman. 1990. Personality Structure: Emergence of the Five-Factor Model. *Annual Review of Psychology* 41, 1 (1990), 417–440.

[3] Janna Lynn Dupree, Richard Devries, Daniel M. Berry, and Edward Lank. 2016. Privacy personas: clustering users via attitudes and behaviors toward security practices. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 5228–5239.

[4] European Commission. 2015. Special Eurobarometer 431 - Data Protection. (2015). http://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_431_en.pdf

[5] Lujun Fang and Kristen LeFevre. 2010. Privacy Wizards for Social Networking Sites. In *Proceedings of the 19th International Conference on World Wide Web (WWW '10)*. ACM, New York, NY, USA, 351–360.

[6] ISO. 2018. ISO 9241-11:2018 Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts. (2018).

[7] ISSRG. [n. d.]. Permis. ([n. d.]). http://sec.cs.kent.ac.uk/permis/

[8] Maritza Johnson, John Karat, Clare-Marie Karat, and Keith Grueneberg. 2010. Usable Policy Template Authoring for Iterative Policy Refinement. In *IEEE International Symposium on Policies for Distributed Systems and Networks*. IEEE Computer Society, [Los Alamitos, Calif.], 18–21.

[9] David Keirsey. 1998. *Please understand me 2*. Prometheus Nemesis Book Company.

[10] Ponnurangam Kumaraguru and Lorrie Cranor. 2005. Privacy indexes: a survey of Westin's studies. (2005). http://repository.cmu.edu/isr/856

[11] Yabing Liu, Krishna P. Gummadi, Balachander Krishnamurthy, and Alan Mislove. 2011. Analyzing Facebook privacy settings: User expectations vs. reality. In *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*. 61–70.

[12] Naresh K. Malhotra, Sung S. Kim, and James Agarwal. 2004. Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information systems research* 15, 4 (2004), 336–355.

[13] Anthony Morton and M. Angela Sasse. 2014. Desperately seeking assurances: Segmenting users by their information-seeking preferences. In *Privacy, Security and Trust (PST), 2014 Twelfth Annual International Conference on*. 102–111.

[14] Isabel Briggs Myers, Mary H. McCaulley, and Robert Most. 1985. *Manual: A guide to the development and use of the Myers-Briggs Type Indicator*. Vol. 1985. Consulting Psychologists Press Palo Alto, CA.

[15] Manuel Rudolph, Denis Feth, Joerg Doerr, and Joerg Spilker. 2016. Requirements Elicitation and Derivation of Security Policy Templates—An Industrial Case Study. In *24th International Requirements Engineering Conference (RE)*. 283–292.

[16] Manuel Rudolph, Denis Feth, and Svenja Polst. 2018. Why Users Ignore Privacy Policies: A Survey and Intention Model for Explaining User Privacy Behavior. In *19th International Conference on Human-Computer Interaction (HCII)*.

[17] Manuel Rudolph and Svenja Polst. 2018. Satisfying and Efficient Privacy Settings. *Mensch und Computer* (2018).

[18] Manuel Rudolph, Svenja Polst, and Joerg Doerr. 2019. Enabling Users to Specify Correct Privacy Requirements. In *International Working Conference on Requirements Engineering: Foundation for Software Quality*. Springer, 39–54.

[19] H. Jeff Smith, Sandra J. Milberg, and Sandra J. Burke. 1996. Information privacy: measuring individuals' concerns about organizational practices. *MIS Quarterly* (1996), 167–196.

[20] Jennifer M. Urban and Chris Jay Hoofnagle. 2014. The Privacy Pragmatic as Privacy Vulnerable. In *Workshop on Privacy Personas and Segmentation*.

[21] Andrzej Uszok, Jeffrey Bradshaw, Renia Jeffers, Niranjan Suri, Patrick Hayes, Maggie Breedy, Larry Bunch, Matt Johnson, Shriniwas Kulkarni, and James Lott. 2003. KAoS policy and domain services: Toward a description-logic approach to policy representation, deconfliction, and enforcement. In *Policies for Distributed Systems and Networks, 2003. Proceedings. POLICY 2003. IEEE 4th International Workshop on*. 93–96.