SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

April 2026, Vol. 69, Iss. 4, 142105:1–142105:16
https://doi.org/10.1007/s11432-025-4594-4

On the user friendliness of password creation policy designs in the wild

Ding WANG[1,2,3], Tongxin WEI[1,2,3] & Zhenduo HOU[1,2,3*]

[1]*College of Cryptology and Cyber Science, Nankai University, Tianjin 300350, China*
[2]*Key Laboratory of Data and Intelligent System Security (Nankai University), Ministry of Education, Tianjin 300350, China*
[3]*Tianjin Key Laboratory of Network and Data Security Technology, Nankai University, Tianjin 300350, China*

Received 28 April 2025/Revised 18 July 2025/Accepted 15 September 2025/Published online 5 March 2026

Abstract Password creation policy designs (PCPDs) are the design of site interfaces for user password registration. Much attention has been paid to the security of password policies, but there is a lack of research on the presentation of PCPDs (i.e., password rule (PR), password registration error message (PREM), and password strength meter (PSM)). To fill this gap, we, for the first time, evaluate the user-friendliness of PCPDs across 163 Chinese and 202 U.S. websites. We define two key criteria for PCPD user-friendliness: friendly timing and friendly explanation. Our results present a concerning picture: only 8.6% (14/163) of Chinese sites and 3.5% (7/202) of U.S. sites meet our user-friendliness criteria for PCPDs. To validate our user-friendly PCPDs, we conduct a survey with 249 U.S. and Chinese participants, confirming that over 75% of participants support our criteria for optimal password creation interfaces. Furthermore, we conduct a correlation analysis of the security of the PCPD and corresponding password datasets from eleven Chinese and U.S. websites. We reveal that an unfriendly PCPD increases the time for password creation and login, but does not improve password security. Our work highlights that leading services need substantial improvements in the usability of PCPDs during the user registration process.

Keywords password creation policy, user-friendly design, password security, online experiment, offline interview

Citation Wang D, Wei T X, Hou Z D. On the user friendliness of password creation policy designs in the wild. Sci China Inf Sci, 2026, 69(4): 142105, https://doi.org/10.1007/s11432-025-4594-4

1 Introduction

Passwords remain the most prevalent authentication method across various online platforms, e.g., social media, e-commerce sites, and other digital services [1,2]. This popularity is due to the cost-effectiveness of password systems, their ease of reset, and simple usage, unlike other authentication methods (e.g., multi-factor authentication or mobile one-time passwords that may require additional tools). Additionally, the cost of switching from passwords to other systems is hard to calculate, and often, the advantages do not seem to outweigh the significant effort and expense required for such a change.

System-assigned passwords are often difficult to memorize by users [3, 4]. In our survey, 90.7% of Chinese users and 82% of American users prefer creating their own passwords over using system-assigned ones (see Section 5). For human-chosen passwords, people tend to choose passwords that are easy to remember and short [5]. However, this choice makes their accounts more vulnerable to hacking [6, 7]. Recent research shows that well-thought-out password creation policies can steer users towards selecting passwords that balance security and ease of recall [8,9]. Generally, a password creation policy design (PCPD) typically involves design elements such as password rules (PR), password registration error messages (PREM), and a password strength meter (PSM). PRs make sure that users create passwords that are complex enough, like those that include a mix of letters and numbers, encouraging the creation of stronger passwords [10]. PSM offers password strength feedback on how robust a password is during the registration process [11]. PREM explains what is wrong if the passwords do not comply and prompts users to revise their passwords [12]. This approach highlights the importance of effective password policies for online security.

As we all know, password security faces ongoing threats from password guessing [13–15] and leaks [16]. To tackle this, researchers have rolled out some key strategies in password creation, like setting up complex password rules [17, 18] that require using different types of characters and applying accurate strength meters [19, 20] to encourage

* Corresponding author (email: joehou13@nankai.edu.cn)

stronger passwords. However, strict password rules required by websites can often leave users frustrated [21]. This complexity in the password registration process can greatly decrease the ease of creating and managing passwords. Though a number of user studies have predominantly focused on the security of password rules [18,22,23] or password strength meters [19, 24], they often overlook how websites present and communicate password creation policies to users. Specifically, the content and presentation of password rules have a significant impact on the effectiveness of users' password creation efforts [5, 9]. Overlooking this aspect can lead to poor presentation of PCPDs, which in turn can complicate the user experience during the password registration process. This highlights the need for clear, concise, and user-friendly PCPDs to enhance security while maintaining usability [25].

1.1 Motivations

Although website password policies are widely deployed across the Internet, how these policies are presented by different websites remains a contentious issue. Some prior research (e.g., [1, 2, 11]) has conducted user studies on PRs, PREMs and PSMs, but they primarily focus on the user-friendliness of sites' PCPDs. Relatively little attention has been given to investigate users' expectations during the password registration process. In 2022, Lee et al. [1] conducted a survey on the deployment of website password policies and found that many sites failed to effectively implement password strength meters, with some even unable to reject common weak passwords, e.g., "abc123456." These findings suggest that even high-value service websites do not surpass the industry norms of regular sites in terms of presenting their password policies. Additionally, Fujita et al. [26] discovered that only 35% of 231 Japanese websites employed user-friendly PREMs, highlighting a greater need for usable security. This reveals a significant inconsistency in the presentation of password policies across websites (e.g., the display form of PRs, PREM and PSM), highlighting the absence of a widely accepted industry standard for password creation policy design. Clearly communicating password requirements, offering real-time PSM, and presenting clear PREMs can motivate users to create more secure passwords [2]. However, completing password creation within the interface poses challenges for users in fully understanding and adhering to the PCPDs.

As far as we know, Fujita et al.'s work [26] at SOUPS'22 may be the closest to our paper: they conducted a survey on PREMs (i.e., interpretability and efficiency). But our work differs in the following five key aspects. (1) We explore three design aspects (i.e., PRs, PREMs, and PSMs) during the password registration process, while SOUPS'22 paper [26] focuses exclusively on the PREM. PREM only appears when the password does not meet the website's rules, while PR and PSM are equally important and cannot be overlooked. (2) Unlike SOUPS'22 [26], which only referred to user-friendly guidelines to define and monitor websites, we invite Chinese and U.S. participants to engage in a real-world simulated password registration process, capturing genuine user reactions and time costs. Our goal is to clarify how different PCPDs influence the usability and security of passwords. (3) To further investigate the impact of PCPD on users' password security, we assess the overall security of leaked website passwords (e.g., password length, entropy, composition, and resistance to attacks) based on different PCPDs, a topic not explored in related work [1, 26]. (4) We find that 56.4% of Chinese websites and 26.24% of U.S. websites use friendly PREMs, which is comparable to the finding (35%) in SOUPS'22. However, when considering all three PCPDs, only 8.6% of Chinese websites and 3.5% of U.S. websites meet all three "friendly" PCPDs. This issue has not been highlighted in previous studies (e.g., [1, 26]). (5) We offer new insights (e.g., unfriendly PCPDs not only waste users' time but also do not improve password security) beyond prior studies [1, 26]. In summary, we try to answer the following key research questions (RQs).

RQ1. [User expectations]: What are users' perceptions of user-friendliness and their expectations regarding the usability of password creation policy designs?

RQ2. [Service practices]: Do Chinese and U.S. sites adopt friendly password creation policy designs?

RQ3. [Usable security]: Do password creation policy designs affect users' password security?

Addressing the aforementioned research questions is essential to facilitate the design, evaluation, and selection of user-friendly guidelines in password creation policy design for any given site. Our work represents the initial step towards tackling these questions.

1.2 Contributions

In this paper, we take the first step by conducting user surveys to understand users' perceptions of user-friendliness and expectations regarding PCPDs, as illustrated in our workflow (see Figure 1). It consists of four main steps: (1) our researchers outline the criteria for friendly PCPDs based on guidelines (see Section 6); (2) we assess the user-friendliness of PCPDs on U.S. and Chinese sites (see Section 3); (3) we conduct a user study using PCPD examples derived from the site surveys (see Section 4). We analyze users' perceptions of user-friendliness and expectations regarding PCPDs to determine if they match our defined criteria (see Section 5); (4) we examine the impact of

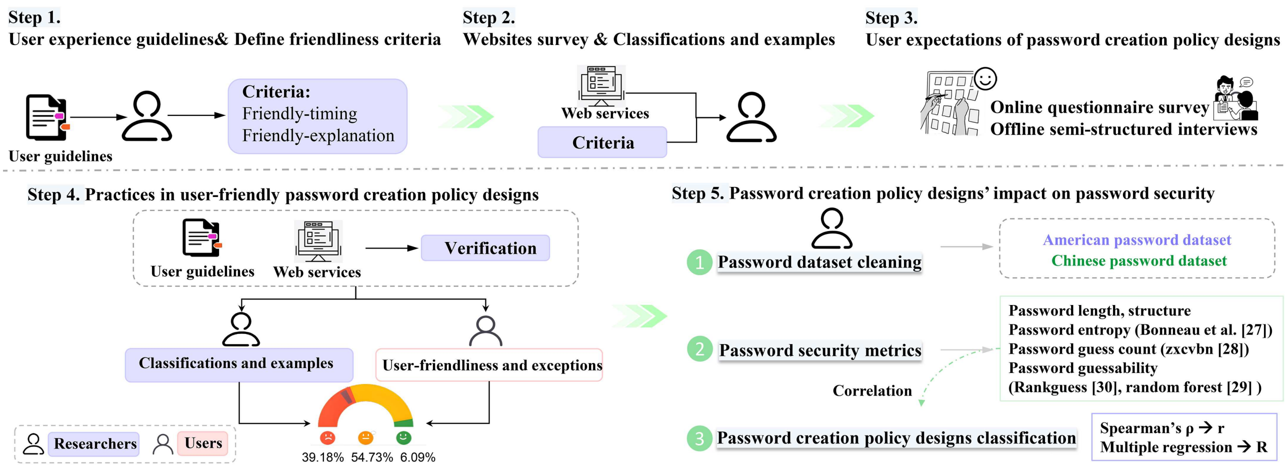


Figure 1 (Color online) Workflow of our study on the user-friendliness of password creation policy designs (PCPDs).

various PCPDs on password security for websites corresponding to eleven leaked password datasets. Most Chinese and U.S. websites have unfriendly PCPDs, which waste users' time during password creation and login without effectively enhancing password security.

The key contributions of our study are four-fold.

- **A new issue.** This study takes the first step toward assessing the user-friendliness of PCPDs from a user perspective, uncovering a significant issue: the majority of sites do not adopt user-friendly PCPDs, and unfriendly PCPDs hinder users' password registration and login processes while failing to enhance password security. We, for the first time, propose a systematic evaluation framework, built on guidelines and criteria across different dimensions, to assess the user-friendliness of PCPDs (e.g., PRs, PREMs and PSMs). Our evaluation framework can be applied to evaluate other user-friendly scenarios with slight modifications. It facilitates systematic evaluation of existing and emerging PCPDs and informs the design of more friendly and secure password creation policies.
- **New sites survey.** Due to the inconsistency in the registration steps across websites, automated tools have difficulty correctly collecting registration information. We manually register on the 163 Chinese sites and the 202 U.S. sites to evaluate the user-friendliness of their PCPDs (see Section 3). We find that only 8.6% (14/163) of Chinese sites and 3.5% (7/202) of U.S. sites meet the user-friendliness criteria (e.g., PR, PREM, and PSM; see Section 6).
- **Impact of PCPDs on password security.** We evaluate and analyze password length, entropy [27], guessability [28], composition, and resistance to guessing [29,30] across 11 publicly available datasets. Our findings show that unfriendly PCPDs do not improve the overall security of users' real passwords on websites (see Section 7.2). Additionally, we use Spearman's ρ and multiple regression to confirm that there is no significant correlation between the user-friendliness of PCPD and password entropy or length.
- **Some insights.** Based on our experiments, we derive several key insights and offer actionable recommendations to advance the overall password ecosystem (see Section 8). We confirm that unfriendly PCPDs not only hinder users' password registration and login processes but also fail to improve password security. We offer various suggestions for password policy makers, system vendors, and users to improve password security and usability.

2 Related work

As introduced in Section 1, a PCPD typically incorporates PR, PREM, and PSM. We review prior research on PR, PREM, and PSM that examines the impact of PCPDs and usable security practices.

2.1 Password creation policy designs' impact on password security

Prior research (e.g., [2,4,10]) has underscored the significant relationship between password policy and the security of users' passwords. Shay et al. [10] surveyed 470 users, revealing that stricter PRs result in longer and more secure passwords. Concerning PREM, Shay et al. [9] investigated the impact of PREM through a study involving 6435 participants, emphasizing the role of PREMs in guiding users toward more secure password choices. In the context of PSM, Ur et al. [31] discussed how PSMs help users in generating more secure passwords. Nonetheless, these studies have neither explored the connection between PCPDs and user experience guidelines, nor have they addressed the specific requirements of PCPDs or the prevalence of their implementation across sites.

2.2 Usable security in password creation policy designs

Recently, there is a growing interest in how effectively PCPDs are implemented on real-world sites (e.g., [1, 26]). Researchers delve into practical applications, striving to enhance the real-life usability of these policies. Lee et al. [1] have highlighted that a significant number of authoritative sites are yet to master the implementation of robust password policies, with 75% of them failing to block weak passwords like "`abc123456`." This research focuses on the practices around password policies, aiming at usable security but frequently misses how PCPDs directly affect the user experience. This shows that there remains a noticeable gap in user-centered research regarding how users perceive the usability of PCPDs.

At SOUPS'22, Fujita et al. [26] took a closer look at the user-friendliness of PREMs across 231 Japanese sites. Their findings indicated that only around 35% of these sites conformed to user-friendly design standards, revealing a widespread discrepancy in how authentication sites meet the practical design guidelines. Despite referencing User Experience guidelines, Fujita et al.'s study lacks direct user engagement to verify these user-friendliness criteria, pointing to a gap between theoretical guidelines and actual user feedback. Their primary reliance on guidelines, without conducting a user study, might not fully substantiate their findings. We propose that incorporating a detailed user survey and evaluation could significantly bolster the credibility of their conclusions. Additionally, their research is limited to Japanese sites, without considering sites of other countries (e.g., U.S. and China). To address this oversight, we gather and categorize sites' PCPDs based on criteria for user-friendliness. We then invite users to evaluate the user-friendliness of these PCPDs and verify the effectiveness of our criteria for user-friendly PCPDs.

2.3 The impact of usable security

For over two decades, researchers have delved into the realm of usable security [32]. Frøkjær et al. [33] argued that efficiency and user-friendliness should be treated as distinct metrics in usability testing. Green et al. [34] stressed the importance of adopting developer-friendly approach to enhance security systems. Golla et al. [2] conducted a study with 180 participants to evaluate the effectiveness of notifications in discouraging password reuse, concluding that notifications alone are inadequate for tackling the problem.

Despite extensive research on password policies, the enhancements to the password registration process have been modest. Recently, a growing number of researchers have dedicated their efforts to improving usable security [35–37]. A few earlier studies [17, 38, 39] have examined the requirements for password rules during registration and how password strength meters influence password security and usability. Over the years, definitions of user-friendliness have been based on user experience guidelines [40, 41]. Although these guidelines set clear benchmarks, some research (e.g., [26, 42]) overlooks capturing the actual expectations and experiences of users in specific contexts. Our user studies address this limitation.

3 Website survey

To tackle RQ1, we evaluate the user-friendliness of PCPDs on a total of 365 sites, including 202 U.S. sites [43] and 163 Chinese sites [44] in February 2023. Since the registration steps and processes vary across websites, automated tools are ineffective on account registration interfaces [26,45]. We manually register on these sites to capture reliable screenshots, develop evaluation criteria based on genuine user experiences, and effectively gauge the usability of PCPDs.

Site URL retrieval. We compile a list of existing site rankings and identify key factors that influence the usability of PCPDs. To provide a thorough overview of the password registration process, we examine twenty web service themes, including e-commerce, accommodation, gaming, information technology, jobs, education, and more. We used the Chinese site ranking service [44] and the U.S. site ranking service [43] for the task conducted from June to July 2023. This approach ensures a comprehensive evaluation of PCPDs across a wide range of online services. In our analysis of the 500 Chinese websites, we successfully access 447 sites, while 53 are inaccessible. Of the accessible sites, 163 employ password registration, 133 allow registration without a password, and 151 do not require identity authentication. For the 500 U.S. websites, we access 319 sites, with 131 being inaccessible. Among these, 202 sites require users to register with a password, 56 offer registration without a password, and 61 do not demand identity authentication (see Table E1 in Supplementary File). We hire seven workers, who are undergraduate computer security students, to register on 163 Chinese and 202 U.S. sites using specific password samples.

Password registration process. We create an account for password registration, which includes a new username, phone number, and email address. We record with screenshots, not limited to the operation steps, conditions of account registration, and any notifications or instructions. These screenshots serve as compelling evidence and examples for subsequent user research (see Table B1 in Supplementary File). We try to identify the critical steps and potential obstacles in creating passwords.

Password samples. To gather PCPD practices, we use password samples to examine how sites respond. Leveraging the Top 10 most popular passwords lists [46] for our empirical analysis, we modify these passwords to assess the PCPDs provided by various sites. We evaluate the sites with "`123456`" (e.g., "`123`", "`12345678`", "`123456abc`", "`123456Abc`", "`123456Abc.`"). Additionally, we include the password "`w2i@K8Duh!91`," which has a guess number exceeding one trillion with zxcvbn [28], to test the effectiveness of the sites' PSM. We capture genuine user experiences and actual PCPDs through manual registrations. Seven workers are undergraduate students with a background in cryptography and have a detailed understanding of the survey's rules (more details can be found in Supplementary File).

Results of site survey. We categorize the display methods of PRs into four types: shown before the user creates the password, shown when the input box is clicked, shown when an error occurs, and not shown at all. The PSM may either be provided or not, with the former including strength indicators and explanations. Password registration error messages include: step-by-step showing errors, displaying all errors, not showing any errors, and showing feedback only after clicking the submit button. We presents the results of PCPDs across the websites (see Section 6) and the PCPD interfaces of these site registrations (see Figures F4–F15 in Supplementary File). Additionally, we summarize the common distribution of components during the registration process. The reasons for the unavailability of other sites include (1) 54.2% have ceased offering services without registration; (2) 9.2% use verification codes sent to the user's mobile phone or email for registration, bypassing traditional password registration; (3) 6.3% do not require registration at all (e.g., *Xinhuanet*, *Globalnet*, *Weathernet*); (4) 2.5% direct users to download a mobile app for registration; (5) 1% are inaccessible due to IP restrictions, including *Yahoo*.

4 User study

The survey questions are designed based on the PCPD framework and users' password creation behaviors. Before the main study, we conduct three pilot studies to refine the phrasing and clarify the distinction between friendly explanation and technical explicability. Each question corresponds to a key aspect of how users evaluate the friendliness of PCPDs. The complete questionnaire is available in Appendix A in Supplementary File. Our research addresses RQ1 through the following steps.

4.1 Online survey

The survey comprises three main parts. In the first part, we introduce participants to the entire password registration process. We ask about their familiarity with the password registration process, excluding those who have never participated in a password registration before participating in the survey. At the same time, we collect data on users' preferred methods for creating passwords. In the second part, we categorize PCPDs into groups according to user-friendliness of definitions (see Section 6) and guide participants to specific URLs for registration (see Table C1 in Supplementary File). Consequently, we introduce two new metrics: registration efficiency and text friendly-explanation, to assess alignment with user experience guidelines (discussed in Section 6). We invite users to rate the PCPDs' user-friendliness from time efficiency and text friendly-explanation. Additionally, to investigate the impact of PCPD on user behavior, we inquire about users' password-related practices across various PCPD scenarios. Finally, we collect demographic information. The questionnaire's content remains the same across different languages.

4.2 Semi-structured interview

To better understand what user-friendly PCPDs are, we conduct interviews with participants (see the details of the offline interview in Appendix B in Supplementary File). The interview process begins with the moderator addressing any questions participants may have about the consent form, which is sent via email prior to the interview. Upon obtaining explicit consent, the recording begins. Participants are reassured that their personal information will

Table 1 Demographics of the questionnaire participants (Chinese participants (N = 96), American participants (N = 153)). CP = Chinese participants; AP = American participants.

Gender			Age			Education			Professional background			Proficiency		
	CP (%)	AP (%)		CP (%)	AP (%)		CP (%)	AP (%)		CP (%)	AP (%)		CP (%)	AP (%)
Male	60	50	18–25	31	20	Below bachelor degree	22	42	Irrelevant	23	33	Not familiar	15	26
Female	40	48	26–35	36	12	Bachelor's degree or equivalent	43	34	Not relevant	9	24	Basic	40	46
No response	–	2	36–45	14	24	Master's degree or equivalent	26	20	Normal	26	25	Familiar	29	23
			46–55	16	26	Doctor degree or equivalent	9	4	Relevant	23	11	Developer/Professional	16	5
			56–65	3	10	No response	–	–	Highly relevant	19	3	No response	–	–
			66+	–	8									

remain strictly confidential and will not be disclosed outside the research team, ensuring adherence to ethical research standards.

Part 1: Experiment. In this section, we provide participants with computers, new email addresses, mobile phone numbers, and fresh user information, setting the stage for them to engage in the password registration process. We categorize the PCPDs into groups based on their user-friendliness, focusing on the efficiency and interpretability of various components in the account registration process. We then introduce our organized PCPDs (i.e., PR, PSM, PREM, and other components) and ask participants to evaluate the user-friendliness. The overall rating of a website depends on two key aspects: efficiency and friendly-explanation, which significantly impact users' experience and their perceptions of the account password registration process. A higher score indicates that the site offers a simpler and more streamlined password creation process. We select some sites for our study (see Table C1 in Supplementary File).

During the experiment, participants are randomly assigned a site from each PCPD category. We instruct them to interact with these accounts as though they were their own. We measure the time taken by users to successfully create a password and count the number of mouse clicks during the password creation process. After completing the interview, users are asked to attempt logging into the account, and the number of failed login attempts is recorded.

Part 2: Interview. After the experiment, we conduct a semi-structured interview to address any concerns participants have about different PCPDs. We encourage participants to share their experiences with password registration and provide feedback on the security, usability, and expectations of PCPDs.

4.3 Recruitment and demographics

Online survey. In our online survey, we recruit 321 participants via online research platforms, 183 from Prolific in the U.S. and 138 from Wenjuanxing in China. Prolific, a crowdsourcing platform, is efficient in data collection [47]. Wenjuanxing, known for its easy operation, allows users to easily navigate the platform and receive effective payment for their participation. We conduct a power analysis using G*Power software (F tests, $\alpha = 0.05$, $f = 0.25$, Power $(1 - \beta) = 0.95$, groups = 5), which indicates that a minimum sample size of 305 is required to ensure statistical significance and reliable research conclusions. Based on the participants' individual preferences in the categorical data and in accordance with our pre-defined friendly criteria, we have decided to cease the distribution of the survey. This signifies that we have gathered sufficient data and achieved our desired objectives. We reject responses completed in less than 3 min to ensure quality and obtain 249 (78%) qualitative answers. The participants comprise 153 in the U.S. and 96 in China. The survey takes place from August 2023 to December 2024. Table 1 shows the demographic information. Considering gender, age, and educational background, we seek balanced recruitment with pre-screened demographic data specifying availability. On average, participants spend 12.68 min completing the survey. Based on the local income, we pay them 10 Chinese Yuan in China and 4 US dollars in the US.

Offline interview. We recruit volunteers for the interviews through campus recruitment in China. We collect demographic information from the participants which includes their gender, age range, educational background, etc. Each interview is semi-structured and typically takes 30 min. After interviewing the 16th participant, we do not obtain more diverse results, so data collection continues until 20 participants are reached, at which point data collection stops. We reward each participant with 30 Chinese Yuan in China. The interview is conducted from October 2024 to January 2025. Sixteen participants are college students aged between 20 and 30, with a gender ratio of 1:1. University students are chosen for their insightful perspectives and familiarity with digital services. Four participants are members of the general public, aged between 35 and 45, all holding a bachelor's degree.

4.4 Data analysis

In our study, we employ a mixed-methods approach that combines quantitative and qualitative methods. In our statistical analysis, we apply different tests to understand the relationship between variables. We use a 5-point Likert scale to measure participants' perceptions of the efficiency and friendly-explanation of PCPDs in the user survey. This scale enables the interpretation of user responses along a sentiment continuum, ranging from negative (scores 1–2) to neutral (score 3) to positive (scores 4–5). We employ the Chi-squared test when expected frequencies are at least 5 and Fisher's Exact Test (FET) otherwise, ensuring statistical significance with a significance level (α) set at 0.05 for all tests. The goal is to determine if there are significant associations between different variables, such as user-friendliness and PCPD characteristics. Additionally, we use FET to compare the viewpoints of Chinese and American participants, aiming to identify any shared opinions on user-friendly PCPDs between the two groups. To ensure the reliability of user ratings, we assess consistency using Cronbach's α, with values above 0.7 indicating acceptable reliability. Our two researchers independently perform inductive coding to analyze the open-ended responses, following the guidelines [48]. The inter-coder agreement between the coders is assessed using Cohen's κ, yielding a value greater than 0.7 (0.83), which indicates a significant level of agreement [49]. We resolve coding conflicts through sufficient group discussions among coders, following the practices employed in previous studies [9, 50]. We adjust the codebook by updating the codes.

4.5 Ethical considerations

Password dataset. Our datasets are publicly accessible on the internet and dark web, and are commonly used in password research [14, 51–53]. Although they contain sensitive data, we take three measures to protect users. (1) We store and process all datasets on offline computers. (2) We report only aggregated statistical data to ensure that individual account information is kept confidential. (3) We delete all intermediate sensitive data (e.g., email addresses) after analysis. While these datasets could be misused by others, our work helps security teams identify which PCPDs are more effective in improving user password strength, rather than solely relying on password strength meters [54–56], ultimately benefiting the public. Since the datasets are publicly available, the results of this study can be reproduced. Our use of these datasets has been reviewed by security and privacy experts.

User study. This study is approved by the Institutional Review Board (IRB) at our institution. We review the Menlo Report for general ethical guidelines and ensure compliance with local laws and regulations. Participants give informed consent before completing the questionnaire and can leave at any time without any penalties. We do not collect personal information, and results are stored on the offline computer. We process data offline and report only aggregated data. Interviews take place in a meeting room with consistent conditions. We offer post-survey support to address any concerns.

5 RQ1: user expectations of PCPDs

In this section, we present the scores and perceptions of participants regarding user-friendliness of PCPDs (see Table 2). We answer the RQ1: *What are user-friendliness and users' expectations regarding the usability of password creation policy designs?* Our findings reveal consistent responses from participants in both the United States and China across most queries.

Our results show that 85% of participants prefer creating their own passwords rather than using system-assigned passphrases. 90.7% of Chinese participants and 82% of Americans prefer to design and create passwords independently, while 48.2% of Chinese and 46% of Americans like password reuse. 9.26% of Chinese and 17% of Americans prefer passwords generated by password managers. 5.56% of Chinese and 11% of Americans prefer system-assigned passphrases.

User-friendliness about password creation policy design. As participants ranked the user-friendliness of the PCPDs, we focus on analyzing the data in PR (Chinese participant: rank1; American participant: rank1), PREM (American participant: rank2; Chinese participant: rank3) and PSM (American participant: rank3; Chinese participant: rank2). In PRs, 89.56% (223/249) of participants agree that "Directly Display PR" is friendly (Q1, Q4, average score: 4.41). 83.53% (208/249) of participants prefer "Fully Display PREM" (Q7). 65.46% (163/249) of participants agree that "Step-by-step Feedback" is friendly (Q8). Additionally, over 80% of participants prefer to be guided by a PSM during password registration (Q15). 59% of participants believe PCPDs (e.g., PR, PSM, PREM) play the most crucial role in determining the user-friendliness of the registration process.

5.1 User-friendliness about password rule

In PR, 91% (87/96) of Chinese participants and 89% (136/153) of American participants think that "Directly display PR" is user-friendly. When the PR is shown directly, participants report high user-friendliness with their overall registration experience, with scores ranging from 4.4 to 5. They rate the friendly-explanation of the PR at 4.3, indicating clear and understandable instructions. 55% (137/249) of participants consider "Enter then Display PR" friendly (Q2). For "Click Submit Button then Display PR", 53% (51/96) of Chinese participants and 14%

Table 2 Participants' ratings for the user-friendliness scores for password creation policy designs ($N = 249$). ■ represents 5-point or 4-point, □ represents 1-point or 2-point. We remove the neutral option 3-point, resulting in our total count not being able to reach 100%. ✓ represents that participants are satisfied with the designs (where Cronbach's $0.7 \leqslant \alpha < 0.8$, acceptable). The bolded value indicates the score of the top-ranked friendliness.

	Condition	Score	Participants	Average score			Friendly rank	Timing	Explanation
				User-friendliness	Efficiency	Friendly-explanation			
Password rule	Directly display	■	90	**4.4**	**4.3**	**4.5**	**1**	✓	✓
		□	6						
	Enter then display	■	55	3.4	3.1	3.6	2		✓
		□	22						
	Click submit button	■	29	2.6	2.1	3.2	3		
		□	57						
Password registration error message	Fully feedback	■	84	**4.5**	**4.7**	**4.2**	**1**	✓	✓
		□	4						
	Step-by-step feedback	■	65	3.5	3.2	3.8	2		✓
		□	20						
	Nothing feedback	■	19	1.9	1.2	2.7	4		
		□	74						
	No explanation	■	22	2.3	1.8	2.9	3		
		□	63						
	No PSM	■	27	2.9	2.4	3.4	5		
		□	42						
	Strong or weak feedback	■	65	3.9	**4.1**	3.7	3	✓	✓
		□	13						
Password strength meter	Text explanation	■	80	**4.2**	**4.1**	**4.3**	**1**	✓	✓
		□	6						
	Allow popular passwords	■	70	3.9	3.8	4.1	2	✓	✓
		□	10						
	Block popular passwords	■	42	3	2.6	3.4	4		
		□	38						

(21/153) of American participants view it as friendly (Q3). 33% of participants claim that clicking the submit button and displaying the PR is accepted.

An unfriendly PR display will lead participants to create longer and more complex passwords, and may even cause them to abandon registration. The proportion of Chinese participants who increase their password length due to the influence of PR gradually increases under the three conditions (e.g., displaying password rules after clicking the button, displaying password rules after entering the password, no display of PR), with values of 12.96%, 14.81%, and 22.22%, respectively. For U.S. participants, the proportion of those who increase password length grows under the three conditions, with values of 11%, 13%, and 22%. Additionally, unfriendly PR may lead to Chinese users (9.26%) and U.S. users (13%) abandoning account registration. Participants describe how the display forms of PR influence their password choices, the comment from P3 is a good summation: *"The PRs are similar from site to site, so I can use a pre-designed password to suit the site's requirements. [..] If my action is denied. I create a longer password to register without PRs, which adds to my memory burden. [..]. Certainly! If the site provides clear and specific PRs, I will adjust my password design accordingly to align with those PRs."* (P3)

Memory burden, too many accounts to manage, and lack of energy to create passwords are important factors leading to password reuse among users. Chinese and U.S. participants identify simplified memory (25.93%, 83%), too many accounts and passwords to manage (24.07%, 62%), and saving effort in designing passwords (20.37%, 51%) as the top three factors contributing to password reuse. U.S. participants note personal habits (25%) and websites not preventing password reuse (17%) as influencing factors, while Chinese participants mention unclear PR (11.11%) as a factor affecting their password reuse behavior.

5.2 User-friendliness about password registration error message

Chinese participants' preference shows no significant difference between "Fully Display PREM" (74/96) and "Step-by-step Feedback" (67/96). In contrast, American participants prefer "Fully Display PREM" (134/153) to "Step-by-step Feedback" (96/153). 95% of participants are more inclined to accept the "Fully Display PREM". When participants receive the "Full PREM" feedback all at once, they report high user-friendliness with their overall registration experience, awarding it a score of 4.5 out of 5. They give the friendly-explanation of the PREM a rating of 4.2, indicating that the instructions are clear and easy to understand. Participants view the registration process as highly efficient, with a top score of 4.7, suggesting that the process is simple and fast. A few participants (21, 22%) believe that the absence of a specific explanation is user-friendly, while some (39, 41%) think that not displaying a PREM is user-friendly, even when something goes wrong. Three participants express low user-friendliness with the "Step-by-step Feedback", highlighting its inconvenience (P5, P11, P17). P14 indicates: *"Displaying all warning messages helps me comprehensively judge and design my password."* (P14)

Absence of PREMs on password rules increases users' password reuse and abandonment of registration. The proportion of password reuse among Chinese participants under the four PREM conditions (i.e., "Step-by-step Feedback","Displaying all PREM", "No explanation", "No feedback") is 25.93%, 37.04%, 37.04%, and 37.04%. For U.S. participants that is 4%, 6%, 20%, and 20%. Under the four PREM conditions, the abandonment rates for registration among Chinese participants are 22.22%, 20.37%, 25.93%, and 35.19%. For U.S. participants, the abandonment rates are 9%, 17%, 24.44%, and 31%.

Our surveys demonstrate that 72.7% of users modify their passwords more than once due to website rules or design during password creation. Chinese users modifying their passwords are as follows: 12.96% for once, 25.93% for twice, 50% for three times, 0% for four times, 3.7% for five times, and 0% for more than five times, with 7.41% unsure. For U.S. users, the proportions are 17% for once, 24% for twice, 30% for three times, 6% for four times, 2% for five times, 7% for more than five times, and 15% unsure.

Our results show that 62.79% of Chinese participants and 64% of U.S. participants believe that frequent password changes in registration can cause memory confusion and increase password forgetting. Chinese users are more likely to report increased security risks (46.51%) and password reuse (39.53%) compared to U.S. users (22% and 31%). Meanwhile, U.S. users are more likely to rely on password managers (9%) or face account recovery difficulties (2%), while no Chinese users reported these issues. In our interview, participants (P8, P11, P19) create disposable passwords to meet PRs for less-used accounts. P15 states that frequent password changes during creation cause confusion and make the password hard to remember. *"I gradually encounter PREMs, which constantly modify my password and confuse my final password. The next time I log in, I have to click the 'forgot password' button."* (P15) In interviews, most participants (13/20) highlight the importance of comprehensive PREM. They recommend that websites display all errors and offer detailed feedback to minimize confusion and reduce attempts. *"The more often I change my password, the more likely I am confused and lost track of the final confirmed password."* (P8)

5.3 User-friendliness about password strength meter

We find 42.17% (105/249) of participants consider a PSM essential for password registration. When participants receive the PSM, they express a high level of user-friendliness with their overall registration experience, giving it a score of 4.2 out of 5. The friendly-explanation of the "text explanations" provided by the PSM receives a rating of 4.2, showing that the instructions are clear and straightforward. Participants do not show differences in their feedback on different forms of PSM. In terms of efficiency, the "Strong or Weak" and "Text Explanation" options both receive similar average scores, at 4.13 and 4.14.

Users believe that password strength meters provide more benefits. When PSMs are provided, both groups find them helpful for improving password creation efficiency (Chinese: 46.3% (44/96), U.S.: 45% (69/153)). Due to some participants choosing 'Don't Know,' the number of people who gave a favorable rating did not exceed 50%. However, the proportion of those who considered PSM to be user-friendly remained the highest. Chinese participants (61.11%, 59/96) are more likely than U.S. participants (28%, 43/153) to believe that PSM enhances security. Without PSM, Chinese participants struggle more to assess password strength (50% vs. 20%) and are more likely to increase password complexity (46.3% vs. 15%), while U.S. participants are more likely to rely on memory or reuse passwords from other accounts (11%). Additionally, Chinese users are more prone to overlook security (42.59%) without PSM, while U.S. users tend to choose simpler passwords (46%).

Compared to U.S. users, Chinese users are more confused by PSM. When only PSMs (e.g., strength bars or colors) are shown, Chinese participants are more likely to increase password complexity (61.11% vs. 37%) but also report higher confusion (29.63% vs. 19%). U.S. participants are slightly more likely to adjust their passwords based on strength feedback (46% vs. 38.89%) but are more prone to ignoring password strength and choosing simple passwords (24% vs. 22.22%). When detailed explanations and guidance are provided, Chinese participants show a significant increase in password complexity (68.52% vs. 43%) but still report more confusion (48.15% vs. 19%). U.S. participants are more likely to adjust passwords (44%) but are more prone to ignoring strength and choosing simple passwords (24%).

Although the password strength meters are inconsistent, users are more willing to believe that their passwords are secure. Five participants express that different sites display different password strength results. They suspect that this inconsistency arises from different measurement criteria. Out of the five participants, two participants express understanding and acceptance of this inconsistency (P8, P20). *"I am confident of my password and hope the site can approve its security."* (P8). Many (16/20) participants prefer the site to provide PSM for a heightened sense of security. However, the remaining three participants state they cannot tolerate it (P4, P12, P14). *"I want sites to provide password strength without blocking my selection (like simple passwords)."* (P12)

When a website refuses the use of popular passwords, compared to U.S. users, Chinese users are more likely to modify popular passwords and rely on password managers to store passwords. When a website refuses the popular passwords, U.S. participants (56%) are more likely than Chinese participants (42.59%) to create unique and stronger passwords. Meanwhile, Chinese participants are more likely to modify popular passwords to make them secure (46.3% vs. 28%) or rely on password managers (53.7% vs. 26%). Few participants from either group give up registration (Chinese: 1.85%, U.S.: 2%), and only a small portion of U.S. participants (4%) refer to other accounts' passwords or provide alternative responses, behaviors not observed among Chinese participants.

PCPD significantly influences users' efficiency in password creation and login, with highly interactive setups like 'Block popular passwords' and 'Step-by-step feedback' taking the longest time and most clicks, while simpler setups like 'No PSM' and 'Text explanation' result in faster and fewer interactions, though these outcomes are relative to the specific PRs (see Figure D2 in Supplementary File).

5.4 Comparing perceptions of friendly PCPDs

Upon evaluating the user-friendliness of PCPDs, participants' judgments exhibit a significant difference among American participants (average 2.72, range 1.2–4.66, Q9), whereas Chinese participants show little differences (average 3.40, range 2.65–4.19, Q9, see Table 2).

Participants with more than 20 passwords show a more distinct scoring difference between different types of PCPDs. More concretely, 33% of participants (fewer than five passwords) show no significant preference between the user-friendliness ratings of PR and PREM, with user-friendliness scores of (PR: 4.4, 3.7, 3.83) and (PREM: 4.03, 4.2, 3.4, 3.23) respectively (Q37). These scores reflect slight variations across different PCPDs. In contrast, the 15% of participants managing more than 20 passwords marked a distinct difference in their assessments. Their ratings show significant differences among (PR: 4.75, 3.63, 2.06) and (PREM: 4.38, 3.31, 2.19, 1.5).

Participants with higher educational qualifications tend to have higher expectations regarding the user-friendliness of PCPDs. Participants with lower educational levels tend to be more accepting of certain practices. 80% of participants with a high school education consider displaying PR after clicking the submission button to be user-friendly. In contrast, approval rates decrease among those with higher education: 56% (23/41) of undergraduates, 32% (8/25) of master's degree holders, and 33% (3/9) of doctoral degree holders. Similarly, tolerance for PREMs that lack detailed information and deny login without providing any message, decreases with higher educational attainment. 71% (15/21) of high school-educated participants consider the PREM friendly (Q7). We find that educational background significantly influences user expectations and perceptions, highlighting the importance of tailoring password creation practices to meet the diverse needs and preferences of users across various educational levels.

6 RQ2: practices in user-friendly PCPDs

In this section, we answer RQ2 and present the user-friendliness PCPDs for the 163 Chinese and 202 U.S. sites. We combine user study results (see Section 5) and guidelines (see Section 6) to define the user-friendliness of PCPDs. Following guidelines (e.g., [41, 57, 58]), we define two criteria for user-friendly PCPDs.

Criteria 1 (friendly-timing). In guidelines [41, 58–60], sites are advised to automatically show important messages close to the relevant fields right away. However, the challenge arises because each site has its own cultural characteristics, standard color scheme, and background style. Our main concern is the immediacy of message display, regardless of font type or background. Based on our user research results (see Section 5), PRs should be displayed before password input, PSM should be shown after input, and a PREM should automatically appear following the submission of an invalid password.

Criteria 2 (friendly-explanation). The guidelines [57,61,62] recommend that message prompts offer solutions to avoid errors and provide actionable advice, particularly for users to input a valid password meeting all criteria. Thus, we suggest that sites should clarify why a password is invalid and provide practical guidance and steps to fulfill the registration criteria. Our second requirement specifies that PRs must detail essential criteria like length and

Table 3 (Color online) Password creation policy design of Chinese and U.S. sites in registration. ■ represents the design is friendly, □ represents the design is unfriendly; ∩ denotes the intersection of two conditions, meaning both are satisfied at the same time. The green font indicates the number of websites with user-friendly PCPDs. The bolded value indicates the number of friendly websites under the current conditions.

		A: Password rule				B: Password registration error message				C: Password strength meter (PSM)					A∩B		A∩C		B∩C		A∩B∩C		
		1: Directly display	2: Enter then display	3: Click submit button then display	4: No password rule	1: Nothing feedback	2: No explanation	3: Fully display	4: Step-by-step display	1: No PSM	2: Strong or weak feedback	3: Text explanation	4: Allow popular passwords	5: Block popular passwords	1: Directly display	3: Fully display	1: Directly display	PSM (2, 3, 4, 5)	3: Fully display	PSM (2, 3, 4, 5)	1: Directly display	3: Fully display	PSM (2, 3, 4, 5)
Chinese sites (N = 163)	Number	**92**	48	20	3	2	26	**46**	89	116	**23**	**24**	**43**	**4**	23		33		16			14	
	%	56.4	29.5	12.3	1.8	1.2	16.0	28.2	54.6	71.2	14.1	14.7	26.4	2.5	14.1		20.3		9.9			8.6	
	■		**92** (56.44%)				**46** (28.22%)					**47(28.83%)**			**23**		**33**		**16**			**14**	
	□		71 (43.56%)				117 (71.78%)					116 (71.17%)			140		130		147			149	
U.S. sites (N = 202)	Number	**53**	74	72	3	1	24	**65**	112	159	**26**	**17**	**13**	**30**	25		12		14			7	
	%	26.2	36.6	35.6	1.5	0.5	11.9	32.2	55.4	78.7	12.9	8.4	6.4	14.9	12.4		5.9		6.9			3.5	
	■		**53 (26.24%)**				**65 (32.18%)**					**43 (21.29%)**			**25**		**12**		**14**			**7**	
	□		149 (73.76%)				137 (67.82%)					159 (78.71%)			177		190		188			195	

character types, PREMs should give clear advice for modifications, and the PSM should offer tangible suggestions for improvement or explain the rationale clearly. Based on our user study findings, we classify a PSM as user-friendly as long as the website provides either visual strength feedback or textual explanations.

Our research reveals that (14/163) Chinese sites and (7/202) U.S. sites fulfill the user-friendliness criteria of PCPDs, including PR, PREM, and PSM (see Table 3). 56.4% of Chinese sites and 26.24% of U.S. sites meet the friendly PR. 28.22% of Chinese sites and 32.18% of U.S. sites satisfy the friendly PREM, while 28.83% of Chinese sites and 21.29% of U.S. sites provide a PSM. We survey other components of sites (see Table E2 in Supplementary File).

Password rule: more than half of the Chinese sites (92, 56.4%) display PR directly, whereas only 53 (26.2%) of U.S. sites follow this practice. Chinese sites commonly present PR upfront, while U.S. sites often inform users about PR either after they enter their passwords (74, 36.6%) or upon clicking the Submit button (72, 35.6%). Interestingly, we find that three Chinese and one U.S. site do not enforce PR, allowing simply passwords (e.g., "`123456`").

Password registration error messages: less than half of Chinese (46, 28.2%) and U.S. (65, 32.2%) sites adopt the Friendly PREM. About half of Chinese (89/163, 54.6%) and U.S. (112/202, 55.4%) sites provide step-by-step PREM on passwords rather than displaying all PREM after a single password attempt. Some sites directly reject users' password choices without offering any rationale (Chinese sites: 26, 16.0%; U.S. sites: 24, 11.9%). More critically, some sites (e.g., 3*dmgame* and *livejournal*) allow simple passwords (e.g., "`123456`") for registration without any warning.

Password strength meters: some Chinese (47, 28.8%) and U.S. (43, 21.3%) sites implement PSM to assess password security. Among these, a portion of Chinese (14.7%) and U.S. (8.4%) sites provide textual advice on enhancing password strength, whereas others simply indicate password strength without any detailed guidance (14.1% of Chinese and 12.9% of U.S. sites). Overall, a minority of the sites (28.83% Chinese, 21.29% U.S.) incorporate PSM, and fewer than half of those that do offer textual explanations. While some Chinese (91.5%, 43/47) and U.S. (30.2%, 13/43) sites consider popular passwords with diverse character types to be strong passwords (e.g., "`123456Abc.`").

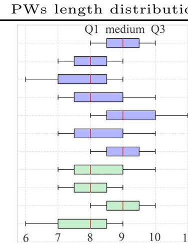
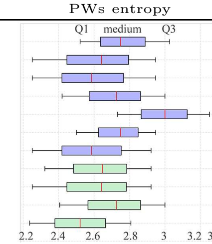
The type of website shows no significant correlation with the user-friendliness of PCPD. Our investigation analyzed 20 types of sites. Regarding PR and PREM, E-commerce, and Retail (PR:8/8; 9/16) and Education (PREM: 8/8; 8/16) demonstrated the highest level of user-friendliness, whereas Games (PR:9/13; 7/15) and Information technology (PREM: 10/13; 8/15) performed poorly. However, neither E-commerce and Retail (PSM:8/8) nor Education (PSM:16/16) provided PSM, whereas types of Games, Information technology, Market research, Music, News and Media, Sports, and Vehicles (PSM:11/12; 9/10; 6/7; 6/6; 12/13; 8/10; 13/14) generally offered PSM.

7 RQ3: PCPD's impact on password security

To analyze the impact of various PCPDs on user password security (RQ3), we conduct a password security evaluation using publicly available leaked password datasets [30, 52] and their corresponding websites. Since users' password behavior is highly sensitive, collecting user behavior data and password change actions often results in volumes of unrealistic password datasets [31]. In our interviews, most users indicate that they deliberately avoided revealing

Table 4 (Color online) Basic information of 11 real-world password datasets. PWs = passwords, PR = Password Rule, PSM = Password Strength Meter, and PREM = Password Registration Error Message. Purple text represents English language web service and green text represents Chinese language web service. **Numbers (1, 2, 3, 4, 5)** correspond to Table 3. Len. = password length. LlDS includes uppercase letters (L), lowercase letters (l), digits (D), and symbols (S). "–" indicates no restrictions on password character types.

Dataset	Web service	When leaked	Total PWs	Len. > 30	Removed	Unique PWs
Wishbone	Social forum	Jan. 2020	10092037	250	0.01%	5933902
Twitter(X)	Social forum	Jun. 2016	40669963	282149	0.69%	10583709
Mate1	Dating website	Mar. 2016	27401505	12430	0.06%	11916080
ClixSense	E-commerce	Jun. 2016	2222045	0	0	1627069
000Webhost	Web hosting	Oct. 2015	15299907	4159	0.76%	10526769
LinkedIn	Job hunting	Jan. 2012	54656615	17162	0.22%	34282741
Yahoo	Job hunting	Jul. 2012	5737798	54224	0.95%	5683574
taobao	E-commerce	Feb. 2016	15072418	88	0.01%	11633759
12306	Train ticketing	Dec. 2014	129303	0	0	117808
CSDN	Programmer forum	Dec. 2011	6428632	0	0.01%	4037605
126	Email	Dec. 2011	6392568	1265	0.024%	3764740

Dataset	PR	PREM	PSM	Len.	PWs length distribution	PWs entropy
Wishbone	4	3	1&5	6		
Twitter(X)	1	4	2	6		
Mate1	1	3	1	6		
ClixSense	2	3	1	8–12		
000Webhost	1	3	1	8		
LinkedIn	1	1	1	6		
Yahoo	2	4	2&5	8		
taobao	1	3	1	6–20		
12306	1	3	2&4	6–30		
CSDN	1	2	1	5–20		
126	1	3	1	8–16		

their actual password behavior, leading to deviations in password security analysis. Therefore, our study focuses on analyzing the authentic websites' overall password security reputation rather than evaluating individual users' password security. Although our password dataset is not extracted under specific environmental conditions, it consists of real leaked user password data, which can still reflect the actual behavior of user passwords.

7.1 Password data and associated websites

Dataset selection. Considering that large-scale collection of real-world password user behavior may raise ethical and privacy risks, we use publicly available leaked password datasets [46, 63, 64] for password security testing. We select seven widely-used U.S. datasets and four Chinese datasets. These datasets cover various languages, service types, and PCPDs, thereby reflecting real-world password characteristics. Although these datasets were leaked 5–14 years ago, they still represent current passwords reasonably well, because password changes over time are marginal [27], and statistical studies on password trends over the last decade have demonstrated that improvements have been rather limited [65]. To ensure consistency, we check the PCPDs of the corresponding websites (via Wayback Machine https://web.archive.org/) prior to the leaks (see Table 4). They collectively contain a total of 1.8 billion passwords. Notably, 000Webhost, a platform primarily used by web administrators, is expected to have a user base with higher security awareness, contributing to a higher level of security in this dataset.

Data cleaning. We discard entries with empty passwords, email addresses missing the '@' character, and incorrectly formatted data (e.g., datasets lacking proper escaping of special characters). Following [21, 24, 63], we exclude strings containing characters beyond the 95 printable ASCII characters. Additionally, we filter out strings with lengths $\geqslant 30$, as manual inspection of the original dataset suggests these longer strings are unlikely to be user-generated. Instead, longer passwords are more likely to be generated by password managers or to be irrelevant data.

Password policy. The Chinese and U.S. websites services in our dataset are various and compositive (see Table 4). In the early stages of password registration, there were no restrictions on character types in password composition requirements and password strength meters. Later, websites upgraded their password policies, requiring at least two types of password characters (e.g., letters and numbers). In password length, Chinese websites enforce maximum length restrictions (16, 20, and 30 characters), whereas U.S. websites generally do not impose such limits. U.S. websites tend to feature more passwords exceeding 30 characters in length (e.g., Twitter). To ensure a fair comparison of password distributions, we restrict our analysis to passwords with lengths of 30 characters or fewer.

7.2 Password security and PCPD relationship

Various factors influence user passwords (e.g., user demographics [33, 66] and password requirements [20, 63]). Thus, we highlight the differences in password characteristics (e.g., password length, structure, entropy and resistance to guessing) between Chinese and American users (see Appendix F in Supplementary File). American users tend to create longer passwords than Chinese users, with higher password entropy. Chinese users prefer numeric passwords

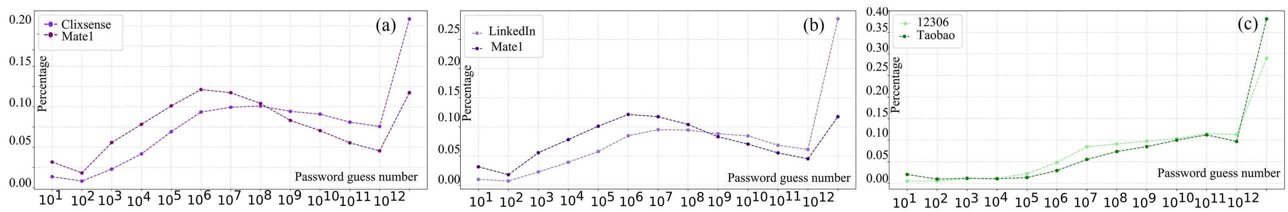


Figure 2 (Color online) Under controlled PR, PREM, and PSM conditions, the percentage of passwords guessed is calculated using Random Forest [29] (training dataset: Rockyou). (a) PR impacts on password security; (b) PREM impacts on password security; (c) PSM impacts on password security. PR stands for password rules, PREM for password registration error messages, and PSM for password strength meter.

and American users favor lowercase-letter passwords. The number of password guesses for Chinese passwords is higher than for English passwords.

To account for these variations, we treat the entire user base of each website as a whole and control for variables to analyze how the website's PCPD impacts the overall security of its password dataset. To control the interference of other influencing factors, we compare password datasets under different PCPDs with the same password requirements (i.e., password length) and language. We use two of the state-of-the-art password guessing algorithms [29,30] to evaluate the online and offline guessing rates of Chinese and English password datasets (training datasets: Rockyou, Dodonew). Using RockYou (English password dataset) as the training set leads to higher cracking rates for English passwords, while using dodonew (Chinese password dataset) results in higher cracking rates for Chinese passwords (see Figure F3 in Supplementary File). Controlling for PCPDs (i.e., PR, PREM and PSM), we assess the correlation between PCPD friendliness and password security using Spearman's method.

Unfriendly PCPDs hinder password creation and login without improving security. In our interviews, participants show low efficiency in password creation and login under unfriendly PCPDs. Figure 2 shows the overall password resistance to guessing across different PCPDs (i.e., PR, PREM and PSM). Although Clixsense and Mate1 show different levels of PR user-friendliness, their password security distributions are strongly correlated ($r = 0.956$, $p < 0.0001$), suggesting a limited effect of PR user-friendliness on password guess resistance. Similarly, while LinkedIn and Mate1 differ in PREM user-friendliness, their password security distributions remain highly correlated ($r = 0.936$, $p < 0.0001$), indicating minimal influence of PREM user-friendliness. Furthermore, despite Taobao and 12306 applying different PSM user-friendliness levels, their password security distributions also show a robust correlation ($r = 0.973$, $p < 0.0001$), suggesting little impact of PSM user-friendliness on password guess resistance. Surprisingly, when the password guess attempts reach 10^6, 12306's (31%) password security is weaker than that of Taobao (25.9%). The password strength evaluation results of zxcvbn [28] and the Random Forest model [29] are consistent on the rockyou dataset, primarily because both rely on similar password feature extraction methods, such as dictionary matching, common pattern recognition, and character combination analysis. We provide more details to verify the impact of PCPD on password guessing resistance (see our Supplementary File). For each subfigure, the two selected sites differ only in the given aspect. For instance, in Figure 2(a), the PRs of Twitter and ClixSense are different (see Table 4), while both employ the same PREM and PSM mechanisms; other subfigures follow the same pattern. In this manner, the closer the two guessing curves are, the less significant the impact of that aspect. We validate the results using other training datasets (e.g., Dodonew), PSM (e.g., zxcvbn [28]), and password guessing RankGuess [30], as shown in Figures F1–F3 in Supplementary File.

8 Discussion and countermeasures

In this section, we summarize key insights derived from our experiments and provide practical suggestions to assist in designing a more user-friendly PCPD. Below we discuss how PCPDs can better assist and protect users in creating usable and secure passwords during the registration process.

The presentation of password policies, along with the policies themselves, significantly affects both user password security and usability. Prior work [1, 22] has primarily focused on password policy (e.g., password length and characters), overlooking the rule presentation methods on websites' password creation interfaces. Our study finds that 85% of participants prefer creating their own passwords, underscoring the need for user-friendly PCPDs (see Section 5.1). However, only 14/163 Chinese websites and 7/202 U.S. websites meet the user-friendly requirements for PCPDs. We confirm that the PCPD user-friendliness of current Chinese and U.S. websites is comparably low, aligning closely with the user-friendliness of Japanese websites, as observed in SOUPS'22 [26].

Establishing and strengthening relevant laws and regulations. Unlike previous user studies [8, 10, 67], which focused more on password security issues caused by password rules, we examine the impact of policy display designs on user behavior and perception. Our findings supplement this by revealing that 72.7% of participants modify their passwords more than once during the creation process (see Section 5.2), which highlights how unfriendly PCPDs can disrupt users' ability to register securely and efficiently. Unfriendly designs not only hinder registration and login but also fail to enhance password security (see Section 7.2). Although there are many standards and rules (e.g., [68, 69]) regarding password security and usability, the management of interface design is still relatively loose. To better align with user preferences and improve security, we urge websites to adopt clear PRs, provide comprehensive PREMs, and utilize effective PSMs in their password creation process.

Websites establish security thresholds for password distribution to ensure a balanced and secure spread of passwords. Websites with password breaches rely on overly simplistic password policies (see Section 7.1), such as setting a minimum password length of just six characters or permitting weak passwords like '123456.' These policies fail to help users create sufficiently secure passwords. The importance of password blacklists has always been emphasized [70], but the scope of the blacklist affects the usability of passwords [71, 72]. We supplement the website survey results on password blacklists, which are absent in Kevin et al.'s work [1]. One-third of leaked passwords are repetitive and common (see Table 4), yet only 4/163 Chinese websites and 30/202 U.S. websites use password blacklists to block these common passwords (see Section 6). 68.52% of Chinese users and 43% of U.S. users believe blocking common passwords improves password security (see Section 5.3). Without PSM, they tend to neglect password security. We urge policymakers to enforce regulations that require websites to block common passwords and implement a site-wide password frequency calculation. When the frequency of repeated passwords across the site reaches a certain security threshold, the system should prevent their reuse, thereby enhancing the overall security of the website's password ecosystem.

Users should collaborate with websites to defend against attacks, rather than adopting an adversarial stance. Users should be educated that password strength requirements aim to enhance account security, not hinder the registration process (see Section 5.3). The website should inform users that account security is a shared responsibility and encourage them to create secure passwords. In our interviews, users express a preference for websites that label their passwords as strong, rather than weak, as this makes them feel confident and reassured (see Section 5). This is consistent with Ur et al.'s observation of users' false sense of security regarding their password safety [64]. Our findings underscore the need to cultivate a proper understanding of password strength in users, rather than leaving it to chance.

Limitations. This study has three limitations. First, a controlled experiment would be an effective way to quantify the causal effects of PR, PREM, and PSM on users' password choices. However, because password creation is a highly sensitive behavior and collecting direct behavioral data raises significant ethical concerns, we rely on publicly leaked password datasets to evaluate large-scale password security under different PCPD conditions. We leave the investigation of password creation by the same users under multiple PCPD conditions as future work. Second, social desirability bias may influence participants' responses. To mitigate this, we inform participants that the study does not evaluate them and encourage honest sharing of their practices. Third, our online survey includes only native English speakers from the U.S. and Chinese participants, and we plan to recruit participants from other countries in future work.

9 Conclusion

We systematically explore how to understand, define, and evaluate the user-friendliness of a site's PCPDs. We, for the first time, study users' perceptions and expectations of friendly PCPDs among 249 participants (153 in the U.S. and 96 in China). This investigation provides valuable insights into the usability of PCPDs and helps define two key user-friendliness criteria. Our user study is based on the PCPDs we proposed, which are summarized from an empirical measurement covering 163 Chinese and 202 U.S. sites. Our findings reveal that only 8.6% of Chinese sites and 3.5% of U.S. sites meet users' expectations. We emphasize the need to enhance users' security awareness and simplify the password creation process. Specifically, we address two open problems left in [26] and one in [73]. Given the multifaceted nature of factors influencing password security, quantitative data alone may not be sufficient to establish direct causal relationships. Unfriendly PCPDs not only hinder users' password creation and login processes but also fail to enhance password security. We believe that our insights will contribute to the development of more user-friendly and secure password creation processes.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant No. 62222208) and Natural Science Foundation of Tianjin, China (Grant No. 25JCJQJC00230).

Supporting information Appendixes A–F. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Lee K, Sjöberg S, Narayanan A. Password policies of most top websites fail to follow best practices. In: Proceedings of USENIX Symp Usable Privacy Secur (SOUPS), 2022. 561–580
2 Golla M, Wei M, Hainline J, et al. "What was that site doing with my Facebook password?": Designing password-reuse notifications. In: Proceedings of ACM SIGSAC Conf Comput Commun Secur (CCS), 2018. 1549–1566
3 Al-Ameen M N, Wright M, Scielzo S. Towards making random passwords memorable: Leveraging users' cognitive ability through multiple cues. In: Proceedings of ACM SIGCHI Conf Human Factors Comput Syst (CHI), 2015. 2315–2324
4 Shay R, Kelley P G, Komanduri S, et al. Correct horse battery staple: Exploring the usability of system-assigned passphrases. In: Proceedings of USENIX Symp Usable Privacy Secur (SOUPS), 2012. 1–20
5 Sadat S E, Lodin H, Ahmadzai N. Highly secure and easy to remember password-based authentication approach. J Soc Sci Util Technol, 2023, 2: 504-515
6 Gulenko I. Improving passwords: Influence of emotions on security behaviour. Inf Manag Comput Secur, 2014, 22: 167–178
7 Pearman S, Thomas J, Naeini P E, et al. Let's go in for a closer look: Observing passwords in their natural habitat. In: Proceedings of ACM SIGSAC Conf Comput Commun Secur (CCS), 2017. 295–310
8 Egelman S, Sotirakopoulos A, Muslukhov I, et al. Does my password go up to eleven?: The impact of password meters on password selection. In: Proceedings of ACM SIGCHI Conf Human Factors Comput Syst (CHI), 2013. 2379–2388
9 Shay R, Bauer L, Christin N, et al. A spoonful of sugar?: The impact of guidance and feedback on password-creation behavior. In: Proceedings of ACM SIGCHI Conf Human Factors Comput Syst (CHI), 2015. 2903–2912
10 Shay R, Komanduri S, Kelley P G, et al. Encountering stronger password requirements: User attitudes and behaviors. In: Proceedings of USENIX Symp Usable Privacy Secur (SOUPS), 2010. 1–20
11 Carnavalet X D C D, Mannan M. A large-scale evaluation of high-impact password strength meters. ACM Trans Inf Syst Secur, 2015, 18: 1–32
12 Jeon H, Petrie C, Cutkosky M R. JATLite: A Java agent infrastructure with message routing. IEEE Internet Comput, 2000, 4: 87–96
13 Jiang J Z, Huang D, Chen H, et al. HTDcr: A job execution framework for high-throughput computing on supercomputers. Sci China Inf Sci, 2024, 67: 112104
14 Pal B, Daniel T, Chatterjee R, et al. Beyond credential stuffing: Password similarity models using neural networks. In: Proceedings of IEEE Symp Secur Privacy, 2019. 417–434
15 Wang D, Zou Y, Xiao Y A, et al. PASS2EDIT: A multi-step generative model for guessing edited passwords. In: Proceedings of USENIX Secur Symp, 2023. 983–1000
16 Michael X. Heiligenstein. Google data breaches: Full timeline through 2022. 2022. https://firewalltimes.com/google-data-breach-timeline/
17 Huaman N, Amft S, Oltrogge M, et al. They would do better if they worked together: The case of interaction problems between password managers and websites. In: Proceedings of IEEE Symp Secur Privacy, 2021. 1367–1381
18 Sahin S, Al Roomi S, Poteat T, et al. Investigating the password policy practices of website administrators. In: Proceedings of IEEE Symp Secur Privacy, 2023. 552–569
19 Golla M, Dürmuth M. On the accuracy of password strength meters. In: Proceedings of ACM SIGSAC Conf Comput Commun Secur (CCS), 2018. 1567–1582
20 Tan J, Bauer L, Christin N, et al. Practical recommendations for stronger, more usable passwords combining minimum-strength, minimum-length, and blocklist requirements. In: Proceedings of ACM SIGSAC Conf Comput Commun Secur (CCS), 2020. 1407–1426
21 Komanduri S, Shay R, Kelley P G, et al. Of passwords and people: Measuring the effect of password-composition policies. In: Proceedings of ACM SIGCHI Conf Human Factors Comput Syst (CHI), 2011, 2595–2604
22 Shay R, Komanduri S, Durity A L, et al. Designing password policies for strength and usability. ACM Trans Inf Syst Secur, 2016, 18: 1–34
23 Weir M, Aggarwal S, Collins M, et al. Testing metrics for password creation policies by attacking large sets of revealed passwords. In: Proceedings of ACM SIGSAC Conf Comput Commun Secur (CCS), 2010. 162–175
24 Wang D, Shan X, Dong Q, et al. No single silver bullet: Measuring the accuracy of password strength meters. In: Proceedings of USENIX Secur Symp, 2023. 947–964
25 Wang C Y, Wang D, Xu G A, et al. Efficient privacy-preserving user authentication scheme with forward secrecy for industry 4.0. Sci China Inf Sci, 2022, 65: 112301
26 Fujita M, Yamanaka T, Matsuda N, et al. Do authentication websites adopt friendly password registration error message design? Proceedings of USENIX Symp Usable Privacy Secur (SOUPS), 2022. 7–9
27 Bonneau J, Herley C, van Oorschot P C, et al. The quest to replace passwords: A framework for comparative evaluation of web authentication schemes. In: Proceedings of IEEE Symp Secur Privacy, 2012. 553–567
28 Wheeler D L. zxcvbn: Low-budget password strength estimation. In: Proceedings of USENIX Secur Symp, 2016. 157–173
29 Wang D, Zou Y, Zhang Z, et al. Password guessing using random forest. In: Proceedings of USENIX Secur Symp, 2023. 965–982
30 Yang T, Wang D. Rankguess: Password guessing using adversarial ranking. In: Proceedings of IEEE Symp Secur Privacy, 2025. 682–700
31 Ur B, Kelley P G, Komanduri S, et al. How does your password measure up? The effect of strength meters on password creation. In: Proceedings of USENIX Secur Symp, 2012. 65–80
32 Payne B D, Edwards W K. A brief introduction to usable security. IEEE Internet Comput, 2008, 12: 13–21
33 Frøkjær E, Hertzum M, Hornbæk K. Measuring usability: Are effectiveness, efficiency, and satisfaction really correlated? In: Proceedings of ACM SIGCHI Conf Human Factors Comput Syst (CHI), 2000. 345–352
34 Green M, Smith M. Developers are not the enemy!: The need for usable security APIs. IEEE Secur Privacy, 2016, 14: 40–46
35 Caputo D D, Pfleeger S L, Sasse M A, et al. Barriers to usable security? Three organizational case studies. IEEE Secur Privacy, 2016, 14: 22–32
36 Gutfleisch M, Klemmer J H, Busch N, et al. How does usable security (not) end up in software products? Results from a qualitative interview study. In: Proceedings of IEEE Symp Secur Privacy, 2022. 893–910
37 Lennartsson M, Kävrestad J, Nohlberg M. Exploring the meaning of usable security: A literature review. Inf Comput Secur, 2021, 29: 647–663
38 de Carnavalet X D C, Mannan M. From very weak to very strong: Analyzing password-strength meters. In: Proceedings of Netw Distrib Syst Secur Symp (NDSS), 2014. 1–19
39 Redmiles E M, Mazurek M L, Dickerson J P. Dancing pigs or externalities?: Measuring the rationality of security decisions. In: Proceedings of ACM Conf Econ Comput (EC), 2018. 215–232
40 Metts M J, Welfle A. Writing is Designing: Words and the User Experience. Brooklyn: Rosenfeld Media, 2020
41 Minhas S. How to Write Good Error Messages. 2018. https://uxplanet.org/how-to-write-good-error-messages-858e4551cd4
42 Golla M, Ho G, Lohmus M, et al. Driving 2fa adoption at scale: Optimizing two-factor authentication notification design patterns. In: Proceedings of USENIX Secur Symp, 2021. 109–126

43 Pochat V L, van Goethem T, Tajalizadehkhoob S, et al. Tranco: A research-oriented top sites ranking hardened against manipulation. In: Proceedings of Netw Distrib Syst Secur Symp (NDSS), 2019. 1–15
44 Xiamen ShareData Technology Co., Ltd. Station Master's Home. 2023. https://top.chinaz.com/
45 Gavazzi A, Williams R, Kirda E, et al. A study of multi-factor and risk-based authentication availability. In: Proceedings of USENIX Secur Symp, 2023. 2043–2060
46 Wang D, Wang P, He D, et al. Birthday, name and bifacial-security: Understanding passwords of Chinese web users. In: Proceedings of USENIX Secur Symp, 2019. 1537–1555
47 Palan S, Schitter C. Prolific. ac—A subject pool for online experiments. J Behav Exp Finance, 2018, 17: 22–27
48 Saldaña J. The Coding Manual for Qualitative Researchers. London: SAGE, 2021. 128–130
49 Habib H, Pearman S, Wang J, et al. "It's a scavenger hunt": Usability of websites' opt-out and data deletion choices. In: Proceedings of ACM SIGCHI Conf Human Factors Comput Syst (CHI), 2020. 1–12
50 Mayer P, Zou Y, Schaub F, et al. "Now I'm a bit angry: " Individuals' awareness, perception, and responses to data breaches that affected them. In: Proceedings of USENIX Secur Symp, 2021. 393–410
51 Pasquini D, Gangwal A, Ateniese G, Improving password guessing via representation learning. In: Proceedings of IEEE Symp Secur Privacy, 2021. 1382–1399
52 Xiu K, Wang D. Pointerguess: Targeted password guessing model using pointer mechanism. In: Proceedings of USENIX Secur Symp, 2024. 5555–5572
53 Wang D, Zhang Z, Wang P, et al. Targeted online password guessing: An underestimated threat. In: Proceedings of ACM SIGSAC Conf Comput Commun Secur (CCS), 2016. 1242–1254
54 Castelluccia C, Dürmuth M, Perito D. Adaptive password-strength meters from Markov models. In: Proceedings of Netw Distrib Syst Secur Symp (NDSS). 2012
55 Melicher W, Ur B, Segreti S M, et al. Fast, lean and accurate: Modeling password guessability using neural networks. In: Proceedings of USENIX Secur Symp, 2017. 1–17
56 Ur B, Segreti S M, Bauer L, et al. Measuring real-world accuracies and biases in modeling password guessability. In: Proceedings of USENIX Secur Symp, 2015. 463–481
57 Cartwright K B. Executive Skills and Reading Comprehension: A Guide for Educators. New York: Guilford Publications. 2023
58 Holden R J, Karsh B T. A review of medical error reporting system design considerations and a proposed cross-level systems research framework. Hum Factors, 2007, 49: 257–276
59 Wu Y, Wang L, Luo J, et al. Upgrading of the EAST distributed synchronization and timing system. IEEE Trans Nucl Sci, 2010, 57: 1522–1527
60 Mohamed C, Ismaili A Z E A. A new leakage optimization approach for IoTs physical design. In: Proceedings of Int Conf Comput Syst Appl (AICCSA), 2021. 1–8
61 Wu B, Campora III J P, Chen S. Learning user friendly type-error messages. In: Proceedings of ACM Program Lang, 2017, 1: 1–29
62 Harding P, Johnson C. Mastering Securities Lending Documentation: A Practical Guide to the Main European and US Master Securities Lending Agreements. London: Pearson UK, 2013
63 Shay R, Komanduri S, Durity A L, et al. Can long passwords be secure and usable? In: Proceedings of ACM SIGCHI Conf Human Factors Comput Syst (CHI), 2014. 2927–2936
64 Ur B, Noma F, Bees J, et al. "I added '!' at the end to make it secure": Observing password creation in the lab. In: Proceedings of USENIX Symp Usable Privacy Secur (SOUPS), 2015. 123–140
65 Blocki J, Liu P. Towards a rigorous statistical analysis of empirical password datasets. In: Proceedings of IEEE Symp Secur Privacy, 2023. 606–625
66 Huang T. Using sor framework to explore the driving factors of older adults smartphone use behavior. Humanit Soc Sci Commun, 2023, 10: 1–16
67 Woo S S, Mirkovic J. Guidedpass: Helping users to create strong and memorable passwords. In: Proceedings of Int Symp Res Attacks Intrusions Defenses (RAID), 2018. 250–270
68 Summers W C, Bosworth E. Password policy: The good, the bad, and the ugly. In: Proceedings of Winter Int Symp Inf Commun Technol (WISICT), 2004. 1–6
69 Khachatryan A. Password policy compliance overview: NIST 800-63, HIPAA, PCI DSS, and the GDPR. 2025. https://www.syteca.com/en/blog/password-policy-compliance-checklist
70 Cheon E, Shin Y, Huh J H, et al. Gesture authentication for smartphones: Evaluation of gesture password selection policies. In: Proceedings of IEEE Symp Secur Privacy, 2020. 249–267
71 Habib H, Colnago J, Melicher W, et al. Password creation in the presence of blacklists. In: Proceedings of Workshop Usable Secur (USEC), 2017. 50–67
72 Segreti S M, Melicher W, Komanduri S, et al. Diversify to survive: Making passwords stronger with adaptive policies. In: Proceedings of USENIX Symp Usable Privacy Secur (SOUPS), 2017. 1–12
73 Greene K K, Choong Y Y. Must I, can I? I don't understand your ambiguous password rules. Inf Comput Secur, 2017, 25: 80–99