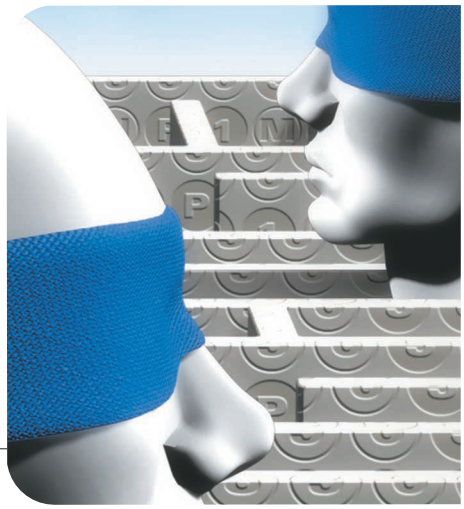
Designing and Evaluating Challenge-Question Systems

The author describes a framework for designing user authentication systems with challenge questions that includes privacy, security, and usability criteria for evaluating a candidate challenge-question system. The proposed challenge-question system for recovering user credentials is based on this framework.



Mike Just
Public Works and Government Services Canada

User authentication is one of the fundamental challenges in the computer-security field and is an enabler for business solutions looking to control access to appropriate individuals. Most people are familiar with authentication in their day-to-day lives—they typically use one or more passwords or identify themselves in some other way, such as over the phone or online. Beyond passwords, other systems can use credentials such as smartcards and biometrics (including fingerprints and retinal scans), offering one or more trade-offs versus password-based solutions.

In this article, I offer a candidate framework for designing challenge-question systems, providing a classification for different question and answer types and discuss how they should meet certain privacy, security, and usability criteria. I also give additional guidance regarding the appropriate number of questions and on the use of complementary security measures. Focus-group testing of a candidate challenge-question system for Canada's Government On-Line (GOL) solution, which I describe in more detail later, provides an initial validation of this framework.

User authentication

Beyond routine authentication, at least two other activities—user identification and recovery—are related to user authentication. Prior to obtaining a password or other credential for subsequent account access, a user is typically identified by the account manager. For example, users can be identified prior to obtaining a password for subsequent access to their bank accounts by demonstrating knowledge of several shared secrets that only the user and the bank would know, such as bank-account number and knowledge of bank-account activity, such as deposits and withdrawals. The password supports a user's further secure access. However, if a user loses or forgets his or her password, there must be a form of user recovery, which could involve reidentification or repetition of the original identification process. However, reidentification is not always possible, because

- *Shared secrets might not exist*. Initial user identification might not have been required—only consistent authentication (such as with an online magazine subscription), in which case the account manager does not know the user's identity. Additionally, privacy concerns could dictate that the account manager not store personal shared secrets (either "at all" or at least not for the purpose of recovery).
- *Shared secrets might not be suitable*. Initial identification might have been done in person and not online. Although in-person registration is acceptable for initial user identification, it could be more cost-effective to support online recovery.

In such situations, when reidentification is neither possible nor preferred as the means of recovery, an initial account registration step includes the addition of some other information that facilitates recovery. At first glance, this might be viewed as cheating—if the other information facilitates recovery, why not use that information for routine authentication as opposed to, say, a password. This can certainly be done, although in most cases, authentication based on other information is more time-consuming than using a password: the quantity of that in-

formation often exceeds that of a typical password to meet the same level of security.

The category of information other than passwords that can be used to authenticate a user is quite vast. For recovery purposes, it is desirable to rely on information the user already knows, rather than requiring him or her to memorize further information; if a user has forgotten a password, he or she might similarly forget additional memorized information. To facilitate consistent presentation of the user's information, organizations typically use challenge questions; to authenticate, the user must correctly answer these questions.

A challenge-question framework

Challenge questions are commonly used as an automated means of password, or more generally, credential recovery. At registration, users select a question for which they also submit a corresponding answer (*answer registration*). During recovery (that is, from a forgotten password), they are challenged with their question and required to provide the appropriate answer (*answer presentation*). (For more on this, see the "Related work" sidebar.)

Clearly, challenge questions offer the same ability to impersonate a user as does credential compromise through other means (such as "shoulder surfing" to observe users entering their passwords). Similarly, the questions offer the same potential for abuse if the system is not usable (for example, when a user writes down the password in a conspicuous location). And if not usable, users might be unwilling or unable to recover, triggering more expensive or manual recovery. In general, all systems that collect user information should be aware of user privacy principles. Thus, the privacy, security, and usability of such systems are a major concern. I describe a framework known as Recovery Question Design and Evaluation Tools (RQDETool), pronounced "are-cue-dee-too," that respects these concerns.

Criteria for challenge question evaluation

As an aid to both the design and evaluation of challenge-question systems, privacy, security, and usability criteria are considered.

Privacy. In environments that use personal information, it should be common practice to follow recognized privacy principles.[1] When using challenge questions and answers to authenticate users, one principle in particular seems relevant: collection limitation, which limits collection of personal user information to only what is necessary for a subsequent recovery process. Although it is important to consider other privacy principles, adherence to this principle helps ensure that only the information necessary to support a suitable level of security and usability is maintained.

Additionally, it would seem natural that the answers to the questions only be used for the purpose of recovering a user's access to his or her account, conforming to a *use limitation* principle. Although not an Organization for Economic Cooperation and Development (OECD) principle per se, it would seem respective of privacy if users had the flexibility and choice to control the personal information, if any, that they provide.

Security. In general, the security of a challenge-question system is concerned with protecting the confidentiality of the answers. (Other properties such as integrity and availability are important to the overall system's security as well, but are not the focus of this framework.) The following security criteria apply primarily to the content of individual questions and answers:

- *Guessing difficulty*. Difficult-to-guess answers would ideally have very high entropy, so that the set of answers are uniformly distributed. In all but a few cases, this is unrealistic; instead, a question that is difficult to guess will simply have a low probability for a successful guess within a small number of attempts. For example, for the question, "What is my eye color?," you would have a high probability for a successful guess (of someone else's eye color) within a small number of attempts because there is a limited set of possible answers.
- *Observation difficulty*. Answers to an unobservable question should be difficult for an attacker to easily retrieve or observe. In particular, the answers should not be available from public sources. For example, for the question, "What is my mother's maiden name?," we would anticipate low to moderate difficulty in discovering the answer for a specifically identified individual (especially because this information is often shared for numerous other applications).

Unlike guessing difficulty, determining a question's observation difficulty is more subjective because the difficulty of determining the answer depends on several factors (for example, the answer's availability). In addition, observation difficulty will differ for individuals that have different relationships with the user—for example, family, friends, acquaintances, colleagues, or strangers.

And if not usable, users might be unwilling or unable to recover, triggering more expensive or manual recovery.

For a recovery system that consists of multiple questions,

Related work on challenge-question systems

Challenge questions are commonly used as an automated means of password or, more generally, credential recovery. At registration, the user selects a question for which he or she also submits a corresponding answer (*answer registration*). During recovery (such as from a forgotten password), the user is challenged with the question and required to provide the appropriate answer (*answer presentation*).

Previous research work

Early work in the area of challenge questions recognized the advantages of using more memorable items for successful user authentication. Although such work performed some valuable usability studies and offered insight into the memorability of answers to challenge questions, insufficient attention was given to ensuring the development of secure questions and for providing a clear framework in which such questions can be designed and evaluated.

Several articles over the past decade or so primarily focused on determining the usability of so-called *cognitive* and *associative passwords*.[1–4] William Haga et al. identified cognitive passwords as questions and answers that relate to the users' facts, opinions, or interests.[1] They further classified them as either *fact-* or *opinion-based* questions. Associative passwords are similar to the game of word association. A word pair is used in which the first word prompts the answer (the second word). For the purpose of a question and answer framework, the first word serves as a question and the second as the answer. Usability studies revealed small variances in the usability of cognitive or associative passwords: where the former were generally easier to remember, they were also more susceptible to guessing by a close family member or colleague.

More recent work focuses on design issues. Carl Ellison et al. examine several issues related to the secure use of challenge questions, and apply secret sharing principles to account for a forgetful user: users must be able to answer $t < n$ questions properly.[5] These techniques also ensure that all questions must be answered simultaneously in order to recover, thus preventing attacks that might attempt to recover individual answers one at a time. Niklas Frykholm and Ari Juels focus on improving usability by designing a system that tolerates the typing mistakes made when entering an answer.[6] Their system uses error-correcting codes, which are helpful for ensuring "repeatability." Lawrence O'Gorman et al. focus on improved user experience, rather than cryptographic issues.[7] They use selectable questions and multiple-choice answers and present several implementations related to this model, which would map to fixed questions and answers in our framework.

Current practice

Challenge questions are used at a variety of Web sites, often in combination with additional protections such as mailing to an address of record (typically an email address). For example, Web-based email providers (such as Yahoo! and Hotmail), and e-commerce sites (such as Amazon, eBay, Chapters, and FutureShop) each use challenge questions in support of account recovery. Online banking services also support a challenge-question system.

From a privacy viewpoint, personal information is sometimes used as part of user identification during recovery for some systems. Most banking sites use personal information as part of account recovery. This is perhaps not too surprising, because the personal information used was directly related to information the banks already had. However, some Web-based email providers will collect additional personal information (such as a date of birth) with the apparent, sole purpose of recovery.

From a security viewpoint, of those solutions in which a user registers a recovery question, only one such question is registered. In most cases, using personal information or mailing to an address of record provides additional security.

References

1. W. Haga and M. Zviran, "Question-and-Answer Passwords: An Empirical Evaluation," *Information Systems*, vol. 16, no. 3, 1991, pp. 335–343.
2. R. Pond et al., "Word Association Computer Passwords: The Effect of Formulation Techniques on Recall and Guessing Rates," *Computers and Security*, vol. 19, no. 7, 2000, pp. 645–656.
3. Y. Spector and J. Ginzberg, "Pass-Sentence—A New Approach to Computer Code," *Computers and Security*, vol. 13, no. 2, 1994, pp. 145–160.
4. M. Zviran and W. Haga, "A Comparison of Password Techniques for Multilevel Authentication Mechanisms," *The Computer J.*, vol. 36, no. 3, 1993, pp. 227–237.
5. C. Ellison et al., "Protecting Secret Keys with Personal Entropy," *J. Future Generation Computer Systems*, vol. 16, no. 4, 2000, pp. 311–318.
6. N. Frykholm and A. Juels, "Error-Tolerant Password Recovery," *Proc. ACM Conf. Computer and Comm. Security* (CCS'01), ACM Press, 2001, pp. 1–9.
7. L. O'Gorman, S. Begga, and J. Bentley, "Call Center Customer Verification by Query-Directed Passwords," *Proc. Financial Cryptography 04*, Int'l Financial Cryptography Assoc., 2004.

we must consider additional criteria, including the total guessing and observation difficulty for the entire set of questions. Additionally, answers should be unrelated so that both their availability and entropy can be independently maintained. One way to support answer independence is to use independent questions (questions that would encourage submission of answers requiring independent attacks).

Usability. In general, a challenge-question system's usability is concerned with providing a user-friendly experience at the stages of answer registration and subsequent answer presentation. When evaluating a challenge-question system, we should use the following usability criteria:

• *Applicability*. The applicability criterion attempts to

characterize the target population's size for which a question might be applicable. For the general population, the question "What was the name of my first pet?," would not apply to those individuals who have never owned a pet. Attempts should be made to support highly applicable questions (though not necessarily at the expense of other criteria).

- *Memorability*. An answer can be considered memorable if the user can easily recall it. This generally implies that the answer should be personally significant. Frequently used information is more memorable, indicating that answers reflecting users' habits, activities, or practices provide suitable answers. For an answer with high recall, only the likelihood of recalling an answer rather than the likelihood of knowing it is considered. For example, although many people might not remember their high-school locker combination, for those who do, it's likely they will be able to continue to recall the answer. However, such a question would only apply to a smaller set of individuals who would know the answer and hence might not be highly applicable.
- *Repeatability*. We must consider at least two aspects of answer repeatability. One, answers should have few syntactic representations. For example, word contractions could cause a discrepancy (such as "St." versus "Street"). Two, answers should also have a semantic value that remains the same over time. For example, the question, "Who is your favorite actor?," is susceptible to the answer changing over time. For this reason, questions asking for "favorites" should be avoided; instead, use "first time" or perhaps "memorable" qualifiers, such as a memorable vacation.

Some additional usability issues will become evident for the specific question and answer types, as I'll discuss next.

Question-and-answer classification

Although other classifications or expansions of my framework might exist, it provides a suitable first step for challenge-question system designers and evaluators (especially in the absence of any other guidance in the literature).

Question types

The two likely most familiar question types are *fixed* and *open*. A fixed question provides the user with a list of preset questions: the choice of question can only be taken "as is" from this list. At the other extreme are open questions, which the user has complete choice of and control over; the system can guide the user as to the question's construction, but the user enters the question in free-form text.

A *controlled question* lies between the extremes of a fixed and open question and consists of a question whose content is partially fixed, although the user can modify it. Two potential variations of a controlled question follow:

- The fixed question might allow the user to add additional text, modifying the original question. The new question would subsequently be presented to the user for credential recovery. For example, to an original fixed question, "What is _____'s favorite food?," the user might add an appropriate name to produce the modified question, "What is Ellen's favorite food?"
- The fixed question could support being combined with an optional user-provided hint, which would be presented to the user as part of credential recovery. For example, to the original fixed question, "What is a memorable date for you?," the user might provide a hint: "Dog." For the user, this hint might indicate a special date associated with their pet dog, such as its date of purchase.

With fixed questions, users don't have to come up with their own questions at registration, which perhaps provides them with an advantage. An open question can potentially improve memorability and applicability for users who can better recognize information that's more memorable to them and construct an appropriate question from this information. However, asking for a completely open question might require too much novelty on behalf of users. A controlled question seems to support a reasonable compromise whereby the user is only responsible for part of the question construction. For example, a challenge question could be as simple as "What is a number that is memorable for you?" (giving the user some content control and guidance); the user can provide the hint, "grade 8 locker," thereby providing advantages similar to an open question. However, controlled questions also share the weaknesses of open questions—the question or hint the user enters can be insecure by providing too much guidance for the answer, making it easier for an attacker. Notice, though, that the hint's repeatability and memorability are not a concern because the user sees the hint upon answer presentation.

With a fixed question, users are prevented from selecting a potentially insecure question such as, "What color are my eyes?," whose answer space is easily exhausted, thereby providing attackers with a security advantage. With an open question, users can select a question that is potentially insecure, although capable users can select more secure questions; for example, they can customize questions directly related and meaningful to their childhood. Additionally, with open questions, users can form associative word pairs, for example, users might associate the word "cat" with the word "my pet," or possibly with "shedding."

When developing questions for a challenge-question system, even further distinctions can be made. One such distinction is that of *fact*- versus *opinion-based* questions.[2] Fact-based questions relate to factual statements regarding a user. We might expect such questions to have less

varying answers over time, and in fact, can be constructed as such, by asking for the first place the user lived rather than his or her most recent residence. However, the answers to such questions might be more readily observable to an attacker because this factual information might be more pervasive. Although opinion-based questions relate to a user's beliefs, and thus might change over time, they should be less pervasive than the answers to fact-based questions (opinions might be less frequently presented and recorded as part of a person's day-to-day activities).

Answer types

A similar distinction applies for fixed, controlled, and open answers as did for their corresponding questions. A fixed-answer set involves user selection of an answer from a predetermined list. At the other extreme, an open answer would involve a user manually entering a response. Answer registration could provide part of the guidance, but the user enters it free-form.

A subtle variation is a controlled answer for which the answer space is not quite fixed or open, but controlled. Some ways in which this is achieved include

- Providing a fixed set of answers for which the answer space is large enough to allow most possible answers. For example, in the case where users answer with a geographic location, they can enter the answer using drop-down menus listing all possible cities, provinces, states, or countries for a region.
- The user can enter an answer, but its format is controlled—answers that don't conform are rejected. For example, in a case in which a user is asked to provide a memorable numeric value, alphabetic and punctuation characters would not be permitted for inclusion in the answer text.

With a fixed-answer set, the system can prevent users from selecting insecure answers. For example, for a given fixed question, there might be a highly probable common answer that the system should disallow or else attackers could easily guess it. However, if there is no unique answer to satisfy a user's preference (either his or her first choice is not available, or more than one satisfactory choice is available), memorability and repeatability might be hampered. Open answers provide larger variation in the answer space, although for certain questions, a user could select highly probable answers. Memorability might be better than with fixed, although repeatability can be a problem if the registered answer is ambiguous (such as in the "St." versus "Street" example). Controlled answers offer an alternative—a large answer space can be used, but control over the possible values improves repeatability, although it doesn't seem to offer any significant security advantages.

An interesting variation to providing the same type of answer at both answer registration and presentation is supported with answers in which the registered answer and that which is presented need not be the same type. For example, registration could allow open answers, in which case the presentation can be fixed. A system could support this option using a set of fake answers that are presented along with the correct answer. Although many further considerations ensure that this option doesn't appear immediately viable, it does illustrate the potential for further variations to our classification.

Question quantity

Once several candidate questions are available, the systems' designers should decide how many questions to use. For security reasons, it is often necessary for the user to register more than one question–answer pair. However, usability tends toward requiring fewer questions. Designers should favor usability over security, so additional security measures are usually enforced thereafter (see the "Complementary techniques" section for more information).

Variations exist in which the number of questions presented at recovery is less than the number of questions registered. There are at least two models:

- The user registers n questions, but is presented with only $t \leq n$ questions on recovery. All t questions must be properly answered for the recovery process to continue.
- The user registers n questions and is presented with $t \leq n$ questions on recovery. Users need only answer $r < t$ questions for the recovery process to continue.

The first option is an attempt to offer a level of security equivalent to that of n questions, while also providing a usability benefit at the time of recovery with fewer questions to the user. However, the usability benefits appear only to reduce the time required for recovery and don't affect the arguably more important concerns of memorability and repeatability (the user still must remember the answers for n questions because he or she won't know what questions will be posed at recovery). Users who register n questions reap some benefit if, after a period of time, they tend to forget the answers to some of these questions as they won't necessarily have to recall and present all n answers. The purpose of the second bullet is to tolerate mistakes upon answer presentation. It seems that an additional question is being used to tolerate such mistakes, but a more usable system might attempt to reduce the number of questions asked.

For a set of candidate questions, some form of "question grouping" might be beneficial. For example, suppose that users must register three questions. It might be advantageous to require them to select one fixed and two controlled questions; for these questions, they must then use a combination of fact- and opinion-based questions. Alter-

natively, questions might be classified based on their topic so that users select one question that requires them to enter a calendar date response, whereas the second might require a numeric response, and so on. If we can classify their questions based on their security strength, then the system could offer multiple classes; the user must then select one question from each class as part of registration.

Complementary techniques

In addition to the questions' construction, evaluation, and grouping, additional techniques can be used for authenticating users, some of which are more suited to a recovery system than for general user authentication. Most notably, mailing to an address of record (typically an email address) is a useful tool. For example, if the address of record is an email address, then the system can email the user an appropriate message giving instructions as part of the recovery process. When combined with a challenge-question recovery system, the email might be sent after the user has successfully answered the challenge questions. Using an address of record provides additional security along with the other security precautions typically in place to control access to the address. Although usability is certainly impacted by adding an additional communication step, the extent to which this is true depends primarily on the amount of time required to complete this step. For some accounts, such latency may not be tolerable.

Additional security measures can improve usability by reducing the number of questions and the security rigour applied to each. This includes a system lockout feature that reduces or removes access to the recovery functionality after several failed attempts. "Graduated lockout" reduces access over time—say, locking out recovery for a fixed period of time after some number of failed recovery attempts—whereas the recovery might be fully blocked after some number of temporary lockouts. Of course, the denial-of-service (DoS) implications must be carefully considered. Reverse Turing tests (for example, the Captcha project; www.captcha.net) help reduce the likelihood of success for online attacks, but this technology is still in its infancy.[3] Alternatively, client puzzles offer a variation for limiting the effectiveness of DoS attacks, whereby the client performs additional computations or "puzzles" before requests can be processed.[4]

A challenge-question system case study

Based on this framework and guidance, a design team with Public Works and Government Services Canada designed a candidate recovery system that is being considered as part of the recovery process for Canada's GOL solution.[5] The GOL initiative's goal is to use information and communication technology to provide Canadians with enhanced access to improved citizen-centered, integrated services, anytime, anywhere, and in the official language of their choice (www.gol-ged.gc.ca/index_e.asp).

Challenge questions

The challenge-question system for the GOL solution is initially constrained; accounts are anonymous, so the central account system neither collects nor retains any identifying information. Therefore, from a privacy viewpoint, using an address of record that can be tied to an individual's identity wasn't acceptable; hence, no out-of-band communication is supported so that a complementary step involving an email to an address of record is not used. Combined with previous focus-group testing in which a five-question system received negative feedback, we designed a system using three questions.

Based on the initial design discussions and focus-group results, fixed-answer solutions were rejected, while open or controlled answers were acceptable. With regard to question type, designers felt that open questions were not sufficiently secure for all citizens. During focus-group evaluation, we were able to obtain some confirmation for this decision because the focus group suggested several insecure questions (see the "Focus-group evaluation and suggestions" section). Therefore, we used a combination of fixed and controlled questions. Based on our collection of a small set of acceptable, fixed questions, we decided to have a single fixed question and two controlled questions. We also attempted to balance between fact- and opinion-based questions.

From these initial considerations, we designed a challenge-question system consisting of

- *Question 1, a set of 15 fixed questions*. The corresponding answer is open. At recovery, the user is presented with a question and asked to enter an answer in free form. Some fixed questions proposed for this list include, "What was my first pet's name?," "Where did I first meet my significant other?," and "What was the last name of my best friend from childhood?" (The focus-group input was used to determine several of these questions.)

Additional security measures can improve usability by reducing the number of questions and the security rigour applied to each.

- *Question 2, a controlled question*. The fixed portion is the question, "Please choose a person who is memorable to you"; the user can also select from a fixed list of hints

that includes "family member," "friend," and "historical figure." The corresponding answer is open, and the user enters it free form. At recovery, the user is presented with the question and an hint, and asked to enter the answer, again in free form.
- *Question 3, a controlled question.* The fixed portion is the question, "Please choose a date that is memorable to you," although the user can also enter a free-form hint. The corresponding answer is controlled, consisting of drop-down selections for the year, month, and day. At recovery, the user is presented the question and a hint, and asked to select an answer from the drop-down dates.

Free-form answers are normalized, removing white space and some punctuation and capitalization. Once users provide answers, a confirmation page confirms the answers.

Focus-group evaluation and suggestions

In August 2003, we led a focus group consisting of 17 individuals from the general population who had experience with the Internet, specifically with online banking or shopping. The group's purpose was twofold:

- to evaluate the proposed challenge-question recovery system; and
- solicit input for new challenge questions.

In general, users accepted the proposed challenge-question system.

Distraction. Some users were distracted by questions that did not apply to them, even though a sufficient number of applicable questions were available; for example, even if there were 10 questions available, and seven or eight were applicable to a user, some users were still distracted by the three or so that weren't applicable. Additional explanatory material (FAQs) could deal with this satisfactorily; not all systems have questions that apply to everyone. To reduce the number of fixed questions that might have to be searched when a user attempts to select his or her preferred question from the fixed list the system to try to use highly applicable questions specific to the given audience.

From a candidate list of approximately 100 questions, we asked participants to rate each with regard to applicability, memorability, and repeatability criteria. In this way, we were able to construct a list of 15 questions as part of question 1, with each rating highly with regard to applicability, memorability, and repeatability.

Memorability. Although questions related to "first time" events are good for repeatability, some older users found them difficult to recall. Therefore, the design should consider various qualifiers, including memorable events. Additionally, the content of the fixed questions should cover events spanning an individual's entire lifetime, thereby more evenly covering a wider age range of users.

Our design settled on using the qualifier "memorable" for questions 2 and 3, with sufficiently broad topics so as to let users customize them by providing a hint.

Uncomfortable with fixed hint. With regard to question 2, some participants were uncomfortable with the choices offered with a fixed hint. In particular, they had difficulty mapping their desired hint to a single selection of a fixed hint. Given the difficulty of providing a hint categorization that a majority of users would approve, the concept of a fixed hint will likely be abandoned (at least for this question).

A subsequent iteration of this system has opted for a free-form hint for question 2.

Remembering dates. Regarding question 3, some participants indicated they could not recall more than a dozen memorable dates, and that most of them are likely observable (that is, either publicly available or known to friends, family, or colleagues, such as birthdays or anniversaries).

Even if a user chooses an observable date, this question should offer protection against a random attack. A targeted attack mounted by someone who knows, or is able to determine, the user's memorable date would still have to answer two remaining questions.

Prefer open questions. Some in the focus group seemed to prefer open questions, although several suggested questions that would not meet a reasonably high security threshold.

A sufficient compromise might be to present an option to either choose from a fixed list of questions or enter a personal, open question (as opposed to just offering an open question). In this way, users wouldn't be forced into possibly constructing a poor question.

Comfortable with multiple questions. Participants were comfortable with the use of three questions. Although many commercial recovery systems rely on only one, other studies have indicated that citizens have higher expectations for security from their government. Therefore, this focus group validated our decision to use three questions.

In addition to the current proposal's evaluation, participants also provided some additional questions to be considered as part of the fixed list for question 1. This exercise provided many creative questions, including, "What was your most elaborate Halloween costume?," "What was your first award?," "What is your secret fear?," "Who is your mentor?," and "Write one word to describe yourself."

Although only one focus group had met at the time this article was written, some validation of our design decisions

was obtained, and we received helpful feedback to adjust those decisions. Further study could involve testing with larger groups and an attempt to determine the memorability after a longer amount of time, such as several months.

Potential future work includes refinements of the challenge-question framework and modifications to the system derived from this framework. The framework will only allow the development of a candidate challenge-question system—for the targeted set of users, usability testing should be performed to confirm the system's suitability.

To put things in perspective, challenge questions offer a form of information for authenticating users. Their main distinguishing feature is the attempt to identify a user based on something known to them (the answer to a question), rather than something that is first memorized (such as a password). Further work in this area should provide other novel solutions, and further validation with users should help to confirm whether a usability advantage is gained. □

References

1. Organization for Economic Cooperation and Development (OECD), "Guidelines on the Protection of Privacy and Transborder Flows of Personal Data," 1980; www.oecd.org.
2. W. Haga and M. Zviran, "Question-and-Answer Passwords: An Empirical Evaluation," *Information Systems*, vol. 16, no. 3, 1991, pp. 335–343.
3. G. Mori and J. Malik, "Up to the Challenge: Computer Scientists Crack a Set of AI-Based Puzzles," *SIAM News*, vol. 35, no. 9, 2002; www.siam.org/siamnews/11-02/gimpy.htm.
4. J. Brainard and A. Jules, "Client Puzzles: A Cryptographic Defense Against Connection Depletion," *Proc. Network and Distributed System Security* (NDSS) *Symp.*, Internet Soc., 1999, pp. 151–165.
5. M. Just, "An Overview of Public Key Certificate Support for Canada's Government On-Line (GOL) Initiative," *Proc. 2nd Annual PKI Research Workshop*, Nat'l Inst. of Standards and Technology (NIST), 2003, pp. 15–26.

Mike Just *is a policy and business strategist in the IT Services Branch of Public Works and Government Services Canada (PWGSC). He is also an adjunct research professor at Carleton University's School of Computer Science. His primary interest is ensuring the delivery of secure yet usable online solutions, both internal and external to government. Prior to his work at PWGSC, he developed IT security policy with the Treasury Board of Canada Secretariat, and worked as an information security specialist at Entrust. He holds a PhD in computer science from Carleton University and is active in the computer-security community. Contact him at www.scs.carleton.ca/~just/.*