How does your password measure up?

The effect of strength meters on password creation

Blase Ur, Patrick Gage Kelley, Saranga Komanduri, Joel Lee, Michael Maass, Michelle L. Mazurek, Timothy Passaro, Richard Shay, Timothy Vidas, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor
Password Meters

• Feedback of password strength at creation time
• Encourage users to create “stronger” passwords
Password Meters

- Widely used

**Google**

Password strength: Weak

Use at least 8 characters. Don’t use a password from another site, or something too obvious like your pet’s name. Why?

**WordPress**

Great passwords use upper and lower case characters, numbers and symbols like !£$%&(*.

**USENIX**

Password strength: Fair

To make your password stronger:
- Add uppercase letters
- Add numbers
- Add punctuation
Password Meters

- Come in all shapes and sizes
Primary Research Questions

1. How do password meters affect ...?
   - Composition
   - Guessability
   - Creation process
   - Memorability
   - User sentiment

2. What elements of meter design are important?

   We present the first large-scale experiment on how the visual and scoring aspects of meters affect password properties
Methodology

• 2,931-participant online study

• Between-subjects design
  • 15 conditions: no meter + 14 different meters

• 2-part study, 2+ days apart
  • Compensated $0.55 and $0.70, respectively
Methodology

• Part 1: Create a password and answer a survey about its creation
  • Scenario: Main email provider is requiring you to change your password
  • Only requirement: 8+ characters

• Part 2: Emailed to return, re-enter password, and answer a survey about remembering password
Overview of Conditions

• **Control** conditions

• **Visual** differences

• **Scoring** differences

• Both **visual & scoring** differences
Control Conditions

- Conditions to which all others were compared:
  - **No meter**: No feedback on strength
  - **Baseline meter**: Standard password meter
Baseline Password Meter

- Based on meters used “in the wild”
  - 8 lowercase letters would fill 1/3\textsuperscript{rd} of the meter

- Higher scores possible from:
  - Longer passwords, e.g., *passwordpassword*
  - More character classes, e.g., *Pa$5wOrD*
Create a password

A strong password helps prevent unauthorized access to your email account.

Type new password: [ ]

8-character minimum; case sensitive

Password strength: Please enter a password in the box above.

Retype new password: [ ]

☑ Make my password expire every 72 days.

Save
Baseline Password Meter

Create a password
Account Password

A strong password helps prevent unauthorized access to your email account.

Type new password: ·······

8-character minimum; case sensitive

Password strength: Poor. Consider adding an uppercase letter or making your password longer.

Non-segmented bar

Retype new password:

Make my password expire every 72 days.

Save
Baseline Password Meter

Create a password
Account Password

A strong password helps prevent unauthorized access to your email account.

Type new password: ******

Password strength: Bad. Consider adding an uppercase letter or making your password longer.

Retype new password: 

Make my password expire every 72 days.

Save

Bad ⇔ Poor ⇔ Fair ⇔ Good ⇔ Excellent
Baseline Password Meter

Create a password
Account Password

A strong password helps prevent unauthorized access to your email account.

Type new password: [Redacted]
8-character minimum; case sensitive
Password strength: Basic
Consider adding an uppercase letter or making your password longer.

Retype new password: [Redacted]

☑ Make my password expire every 72 days.
Save

Provides a suggestion
Baseline Password Meter

Checks against OpenWall cracking dictionary
Baseline Password Meter

Create a password
Account Password
A strong password helps prevent unauthorized access to your email account.

Type new password: 

8-character minimum; case sensitive
Your password is in our dictionary of common passwords.
Password strength: Poor. Consider adding an uppercase letter or making your password longer.

Retype new password:

- Make my password expire every 72 days.

Save

Changes color: red ⇒ orange/yellow ⇒ green
Create a password

A strong password helps prevent unauthorized access to your email account.

Type new password: *********
8-character minimum; case sensitive

Password strength: Good. Consider adding a digit or making your password longer.

Retype new password:

☐ Make my password expire every 72 days.

Save
Create a password

Account Password

A strong password helps prevent unauthorized access to your email account.

Type new password: **********

8-character minimum; case sensitive

Password strength: Excellent!

Retype new password: 

☐ Make my password expire every 72 days.

Save
Overview of Conditions

• **Control** conditions (2)

• **Visual** differences (7)

• **Scoring** differences (4)

• Both **visual & scoring** differences (2)
## Conditions with Visual Differences

<table>
<thead>
<tr>
<th>Condition</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline meter</td>
<td>Bad. Consider adding a digit or making your password longer.</td>
</tr>
<tr>
<td>Three-segment</td>
<td>Bad. Consider adding a digit or making your password longer.</td>
</tr>
<tr>
<td>Green</td>
<td>Bad. Consider adding a digit or making your password longer.</td>
</tr>
<tr>
<td>Tiny</td>
<td>Bad. Consider adding a digit or making your password longer.</td>
</tr>
<tr>
<td>Huge</td>
<td>Bad. Consider adding a digit or making your password longer.</td>
</tr>
<tr>
<td>No suggestions</td>
<td>Bad.</td>
</tr>
<tr>
<td>Text-only</td>
<td>Bad. Consider adding a digit or making your password longer.</td>
</tr>
</tbody>
</table>
## Conditions with Visual Differences

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type new password:</td>
<td>use</td>
</tr>
<tr>
<td>8-character minimum; case sensitive</td>
<td></td>
</tr>
<tr>
<td><strong>Baseline meter</strong></td>
<td>Bad. Consider adding a digit or making your password longer.</td>
</tr>
<tr>
<td><strong>Three-segment</strong></td>
<td>Bad. Consider adding a digit or making your password longer.</td>
</tr>
<tr>
<td><strong>Green</strong></td>
<td>Bad. Consider adding a digit or making your password longer.</td>
</tr>
<tr>
<td><strong>Tiny</strong></td>
<td>Bad. Consider adding a digit or making your password longer.</td>
</tr>
<tr>
<td><strong>Huge</strong></td>
<td>Bad. Consider adding a digit or making your password longer.</td>
</tr>
<tr>
<td><strong>No suggestions</strong></td>
<td>Bad.</td>
</tr>
<tr>
<td><strong>Text-only</strong></td>
<td>Bad. Consider adding a digit or making your password longer.</td>
</tr>
</tbody>
</table>
# Conditions with Visual Differences

<table>
<thead>
<tr>
<th>Condition</th>
<th>Status</th>
<th>Suggestion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline meter</td>
<td>Bad. Consider adding a digit or making your password longer.</td>
<td><img src="image" alt="Red Bar" /></td>
</tr>
<tr>
<td>Three-segment</td>
<td>Bad. Consider adding a digit or making your password longer.</td>
<td><img src="image" alt="Orange Bar" /></td>
</tr>
<tr>
<td>Green</td>
<td>Bad. Consider adding a digit or making your password longer.</td>
<td><img src="image" alt="Green Bar" /></td>
</tr>
<tr>
<td>Tiny</td>
<td>Bad. Consider adding a digit or making your password longer.</td>
<td><img src="image" alt="Red Bar" /></td>
</tr>
<tr>
<td>Huge</td>
<td>Bad. Consider adding a digit or making your password longer.</td>
<td><img src="image" alt="Red Bar" /></td>
</tr>
<tr>
<td>No suggestions</td>
<td>Bad.</td>
<td><img src="image" alt="Red Bar" /></td>
</tr>
<tr>
<td>Text-only</td>
<td>Bad. Consider adding a digit or making your password longer.</td>
<td><img src="image" alt="Red Bar" /></td>
</tr>
</tbody>
</table>

Type new password: [user](image)  
8-character minimum; case sensitive
Conditions with Visual Differences

Type new password: [usenIX]

8-character minimum; case sensitive

Baseline meter: Fair. Consider adding a digit or making your password longer.

Three-segment: Fair. Consider adding a digit or making your password longer.

Green: Fair. Consider adding a digit or making your password longer.

Tiny: Fair. Consider adding a digit or making your password longer.

Huge: Fair. Consider adding a digit or making your password longer.

No suggestions: Fair.

Text-only: Fair. Consider adding a digit or making your password longer.
Conditions with Visual Differences

Type new password: usen1XS

8-character minimum; case sensitive

Baseline meter

Good. Consider adding a digit or making your password longer.

Three-segment

Green

Good. Consider adding a digit or making your password longer.

Tiny

Huge

Good. Consider adding a digit or making your password longer.

No suggestions

Text-only

Good. Consider adding a digit or making your password longer.
Conditions with Visual Differences

Type new password: usen!X$e5!

8-character minimum; case sensitive

Baseline meter: Excellent!

Three-segment: Excellent!

Green: Excellent!

Tiny: Excellent!

Huge: Excellent!

No suggestions: Excellent!

Text-only: Excellent!
A strong password helps prevent unauthorized access to your email account. The stronger your password, the faster Bugs Bunny dances!

- **Type new password:**
- **8-character minimum; case sensitive**
- **Password strength:** Please enter a password in the box above.
- **Retype new password:**
- **Make my password expire every 72 days.**

[Save]
Conditions with Scoring Differences

Type new password: useniX

8-character minimum; case sensitive

Baseline meter

Fair. Consider adding a digit or making your password longer.

Half-score

Bad. Consider adding a digit or making your password longer.

One-third-score

Bad. Consider adding a digit or making your password longer.

Nudge-16

Bad. Consider making your password longer.

Nudge-Comp8

Fair. Consider adding a digit or making your password longer.
<table>
<thead>
<tr>
<th>Conditions with Scoring Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type new password:</td>
</tr>
<tr>
<td>[usenIX$e5]</td>
</tr>
<tr>
<td>8-character minimum; case sensitive</td>
</tr>
<tr>
<td><strong>Baseline meter</strong></td>
</tr>
<tr>
<td>Excellent!</td>
</tr>
<tr>
<td><strong>Half-score</strong></td>
</tr>
<tr>
<td>Poor. Consider adding a different symbol or making your password longer.</td>
</tr>
<tr>
<td><strong>One-third-score</strong></td>
</tr>
<tr>
<td>Bad. Consider adding a different symbol or making your password longer.</td>
</tr>
<tr>
<td><strong>Nudge-16</strong></td>
</tr>
<tr>
<td>Poor. Consider making your password longer.</td>
</tr>
<tr>
<td><strong>Nudge-Comp8</strong></td>
</tr>
<tr>
<td>Excellent!</td>
</tr>
</tbody>
</table>
Conditions with Scoring Differences

Type new password: `usenIX$e5WHYis`

8-character minimum; case sensitive

Baseline meter: Excellent!

Half-score: Fair. Consider adding a different symbol or making your password longer.

One-third-score: Poor. Consider adding a different symbol or making your password longer.

Nudge-16: Good. Consider making your password longer.

Nudge-Comp8: Excellent!
Conditions with Scoring Differences

Type new password: usenIX$e5WHYismyP4$$
8-character minimum; case sensitive

Baseline meter
Excellent!

Half-score
Good. Consider adding a different symbol or making your password longer.

One-third-score
Poor. Consider adding a different symbol or making your password longer.

Nudge-16
Excellent.

Nudge-Comp8
Excellent!
Conditions with Scoring Differences

Type new password: useniX$e5WHYismyP4$$word99
8-character minimum; case sensitive

Baseline meter
Excellent!

Half-score
Excellent!

One-third-score
Fair. Consider adding a different symbol or making your password longer.

Nudge-16
Excellent.

Nudge-Comp8
Excellent!
Conditions with Scoring Differences

Type new password: `usenlX$e5WHYismyP4$$word99notGOOD`

- **Baseline meter**: Excellent!
- **Half-score**: Excellent!
- **One-third-score**: Fair. Consider making your password longer.
- **Nudge-16**: Excellent.
- **Nudge-Comp8**: Excellent!
Conditions with Scoring Differences

Type new password: useniX$e5WHYismyP4$$word99notGOODenough?
8-character minimum; case sensitive

Baseline meter
Excellent!

Half-score
Excellent!

One-third-score
Excellent!

Nudge-16
Excellent!

Nudge-Comp8
Excellent!
Conditions with Visual & Scoring Diff.

Type new password: usenIX
8-character minimum; case sensitive

Baseline meter
Fair. Consider adding a digit or making your password longer.

Text-only-half  Bad. Consider adding a digit or making your password longer.

Bold text-only-half  Bad.
Conditions with Visual & Scoring Diff.

Type new password: useIX$e5

8-character minimum; case sensitive

Baseline meter

Excellent!

Text-only-half

Poor. Consider adding a different digit or making your password longer.

Bold text-only-half

Poor. Consider adding a different digit or making your password longer.
Conditions with Visual & Scoring Diff.

Type new password: usenIX$e5WHY
8-character minimum; case sensitive

Baseline meter
Excellent!

Text-only-half
Fair. Consider adding a different symbol or making your password longer.

Bold text-only-half
Fair. Consider adding a different symbol or making your password longer.
Conditions with Visual & Scoring Diff.

Type new password: usenIX$e5WHYismy|

8-character minimum; case sensitive

Baseline meter: Excellent!

Text-only-half: Fair. Consider adding a different digit or making your password longer.

Bold text-only-half: Fair. Consider adding a different digit or making your password longer.
Conditions with Visual & Scoring Diff.

Type new password: usenIX$e5WHYismyP4$$
8-character minimum; case sensitive

Baseline meter
Excellent!

Text-only-half
Good. Consider adding a different symbol or making your password longer.

Bold text-only-half
Good. Consider adding a different symbol or making your password longer.
Conditions with Visual & Scoring Diff.

Type new password: usenIX$e5WHYismyP4$$word99

8-character minimum; case sensitive

Baseline meter

Text-only-half

Bold text-only-half

Excellent!
Stringent Meters

• Those with more *stringent* scoring:
  • Half-score
  • One-third-score
  • Text-only-half
  • Bold text-only-half
Participants

• 2,931 participants
  • Recruited on Amazon’s Mechanical Turk

• 63% male, 37% female

• 40% technical (degree/job)

• 18 to 74 years old

• 96 countries
  • 42% from India
  • 32% from U.S.
Metrics for Results

- Composition
- Guessability
- Creation process
- Memorability
- Sentiment
Metric – Password Composition

- Password length
Password Length

1: No meter
2: Baseline meter
3: Three-segment
4: Green
5: Tiny
6: Huge
7: No suggestion
8: Text-only
9: Bunny
10: Half-score
11: One-third-score
12: Nudge-16
13: Nudge-comp8
14: Text-only half-score
15: Bold text-only half-score

Condition

Password length
Meters lead to longer passwords
Metric – Guessability

• Guessability – how many guesses to crack?
  • Threat model: offline attack
    • Brute-force first guesses: aaaaaaaaa, aaaaaaab
    • Better first guesses: 12345678, password

• Cracking algorithm: Weir et al. (Oakland ‘09)

• Guess calculator: Kelley et al. (Oakland ‘12)

• Weak adversary: 500 million guesses (5×10^8)
• Medium adversary: 50 billion guesses (5×10^{10})
• Strong adversary: 5 trillion guesses (5×10^{12})
Results – Guessability (No meter)

<table>
<thead>
<tr>
<th>Number of Guesses</th>
<th>Percentage of Passwords Cracked</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0%</td>
</tr>
<tr>
<td>4</td>
<td>10%</td>
</tr>
<tr>
<td>10</td>
<td>20%</td>
</tr>
<tr>
<td>5</td>
<td>30%</td>
</tr>
<tr>
<td>10</td>
<td>40%</td>
</tr>
<tr>
<td>6</td>
<td>50%</td>
</tr>
<tr>
<td>10</td>
<td>60%</td>
</tr>
<tr>
<td>7</td>
<td>70%</td>
</tr>
<tr>
<td>10</td>
<td>80%</td>
</tr>
<tr>
<td>8</td>
<td>90%</td>
</tr>
<tr>
<td>10</td>
<td>100%</td>
</tr>
</tbody>
</table>

- Weak: $5 \times 10^8$
- Medium: $5 \times 10^{10}$
- Strong: $5 \times 10^{12}$

Better

No meter
Results – Guessability (Controls)

<table>
<thead>
<tr>
<th>Number of Guesses</th>
<th>Percentage of Passwords Cracked</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0%</td>
</tr>
<tr>
<td>20</td>
<td>10%</td>
</tr>
<tr>
<td>30</td>
<td>20%</td>
</tr>
<tr>
<td>40</td>
<td>30%</td>
</tr>
<tr>
<td>50</td>
<td>40%</td>
</tr>
<tr>
<td>60</td>
<td>50%</td>
</tr>
<tr>
<td>70</td>
<td>60%</td>
</tr>
<tr>
<td>80</td>
<td>70%</td>
</tr>
<tr>
<td>90</td>
<td>80%</td>
</tr>
<tr>
<td>100</td>
<td>90%</td>
</tr>
</tbody>
</table>

- **Weak**: $5 \times 10^8$
- **Medium**: $5 \times 10^{10}$
- **Strong**: $5 \times 10^{12}$

No meter
Baseline meter
Results – Guessability (Visual)

Number of Guesses

Percentage of Passwords Cracked

Weak
5 × 10^8

Medium
5 × 10^10

Strong
5 × 10^12

No meter
Baseline meter

Text-only
Green
Tiny
Huge
Bunny
No suggestions
Three-segment
Visual changes don’t significantly increase resistance to guessing.
Results – Guessability (Significant)

<table>
<thead>
<tr>
<th>Percentage of Passwords Cracked</th>
<th>Number of Guesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>10^4</td>
</tr>
<tr>
<td>10%</td>
<td>10^5</td>
</tr>
<tr>
<td>20%</td>
<td>10^6</td>
</tr>
<tr>
<td>30%</td>
<td>10^7</td>
</tr>
<tr>
<td>40%</td>
<td>10^8</td>
</tr>
<tr>
<td>50%</td>
<td>10^9</td>
</tr>
<tr>
<td>60%</td>
<td>10^10</td>
</tr>
<tr>
<td>70%</td>
<td>10^11</td>
</tr>
<tr>
<td>80%</td>
<td>10^12</td>
</tr>
<tr>
<td>90%</td>
<td>10^13</td>
</tr>
</tbody>
</table>

- Weak: 5 × 10^8
- Medium: 5 × 10^10
- Strong: 5 × 10^12

No meter

One-third-score

Half-score
Results – Guessability (Significant)

Stringent meters with visual bars increase resistance to guessing.
Creation Process

• Time to create password

• Changed their mind during creation
Results – Password Time of Creation

Time of creation (s)

Condition

1: No meter
2: Baseline meter
3: Three-segment
4: Green
5: Tiny
6: Huge
7: No suggestion
8: Text-only
9: Bunny
10: Half-score
11: One-third-score
12: Nudge-16
13: Nudge-comp8
14: Text-only half-score
15: Bold text-only half-score
Results – Password Time of Creation

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls</td>
<td>No meter</td>
</tr>
<tr>
<td>Visual</td>
<td>Baseline meter</td>
</tr>
<tr>
<td>Scoring</td>
<td>Three-segment</td>
</tr>
<tr>
<td>Visual &amp;</td>
<td>Green</td>
</tr>
<tr>
<td>scoring</td>
<td>Tiny</td>
</tr>
<tr>
<td></td>
<td>Huge</td>
</tr>
<tr>
<td></td>
<td>No suggestion</td>
</tr>
<tr>
<td>10: Half-score</td>
<td>Text-only</td>
</tr>
<tr>
<td>11: One-third-score</td>
<td>Text-only half-score</td>
</tr>
</tbody>
</table>

Time of creation (s)

Condition 1-15
More time was spent creating a password with stringent meters.
Results – Changed Their Mind

• Changed their mind during password creation
  • Typed 8+ character password
    - usenixsec
  • Completely erased it
  • Saved a different password
    - iLIKEbunnies
Results – Changed Their Mind

% who changed mind

Conditions:
1: No meter
2: Baseline meter
3: Three-segment
4: Green
5: Tiny
6: Huge
7: No suggestion
8: Text-only
9: Bunny
10: Half-score
11: One-third-score
12: Nudge-16
13: Nudge-comp8
14: Text-only half-score
15: Bold text-only half-score

Condition

Controls | Visual | Scoring | Visual & scoring

% who changed mind
Results – Changed Their Mind

% who changed mind

<table>
<thead>
<tr>
<th>Condition</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No meter</td>
</tr>
<tr>
<td>2</td>
<td>Baseline meter</td>
</tr>
<tr>
<td>3</td>
<td>Three-segment</td>
</tr>
<tr>
<td>4</td>
<td>Green</td>
</tr>
<tr>
<td>5</td>
<td>Tiny</td>
</tr>
<tr>
<td>6</td>
<td>Huge</td>
</tr>
<tr>
<td>7</td>
<td>No suggestion</td>
</tr>
<tr>
<td>8</td>
<td>Text-only</td>
</tr>
<tr>
<td>9</td>
<td>Bunny</td>
</tr>
<tr>
<td>10</td>
<td>Half-score</td>
</tr>
<tr>
<td>11</td>
<td>One-third-score</td>
</tr>
<tr>
<td>12</td>
<td>Nudge-16</td>
</tr>
<tr>
<td>13</td>
<td>Nudge-comp8</td>
</tr>
<tr>
<td>14</td>
<td>Text-only half-score</td>
</tr>
<tr>
<td>15</td>
<td>Bold text-only half-score</td>
</tr>
</tbody>
</table>
Results – Changed Their Mind

Meters lead people to change their mind during password creation.
Metric – Memorability

• Successful login about 5 minutes later and 2+ days later

• Return rate

• Wrote password down, stored it, or pasted it

• No significant differences across conditions for any of these metrics
Metric – Sentiment

- Level of agreement with 14 statements
  - Password creation (e.g., fun, difficult, annoying)
  - Password meter (e.g., incorrect, high score important)

6. The password strength meter was annoying.*
Results – Sentiment

- Stringent meters a bit more annoying
- Stringent meters violated expectations
The Password Strength Meter Was Annoying

Control
- baseline meter
- three-segment green
- tiny huge
- no suggestions
- text-only bunny

Visual
- half-score
- one-third-score
- nudge-16
- nudge-comp8

Scoring
- text-only half-score
- bold text-only half-score

Visual & Scoring
- Strongly Agree
- Agree
- Neutral
- Disagree
- Strongly Disagree
The Password Strength Meter Was Annoying

Control

Visual

Scoring

Visual & Scoring

Baseline meter: 13%

Three-segment green tiny huge no suggestions text-only bunny

Half-score one-third-score: 27% 37%

Nudge-16 nudge-comp8

Text-only half-score bold text-only half-score: 34% 40%

Bar chart showing the percentage of users who agreed or disagreed with each option.
Takeaway: Meters Matter

- Meters lead to longer passwords
- Stringent meters reduce guessability
- Meters don’t affect memorability
- Overly stringent meters don’t add benefits
  - May even backfire
Takeaway: What Makes Meters Matter

• Important features
  • Scoring stringency
  • Having a visual component

• Less important features
  • Color
  • Segmentation
  • Size
  • Bunnyness
Takeaway: What Makes Meters Matter

• Important features
  • Scoring stringency
  • Having a visual component

• Less important features
  • Color
  • Segmentation
  • Size
  • Bunnyness

Thank You!

Blase Ur (bur@cmu.edu), Patrick Gage Kelley, Saranga Komandur, Joel Lee, Michael Maass, Michelle L. Mazurek, Timothy Passaro, Richard Shay, Timothy Vidas, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor