Testing Computer-Aided Mnemonics and Feedback for Fast Memorization of High-Value Secrets

Sauvik Das  
Carnegie Mellon University

Jason Hong  
Carnegie Mellon University

Stuart Schechter  
Microsoft Research

USEC 2016
Summary

**Question:** Can computer-assisted mnemonics help people learn strong, randomly assigned passwords quickly and with high long-term recall?

**Method:** Design and implementation of two computer-assisted mnemonic training regimens and an experimental evaluation of their effectiveness.
Summary

**Outcome:** Our *story* mnemonic, in which users weave chunks of their random secret into sentences, works really well (7.5 learning sessions, 84% two-week recall)

A pink _ _ _ _ is chasing a giant _ _ _ _ blowing in the _ _ _ _.

cat  leaf  wind
A pink **cat** _ _ is chasing a giant _ _ _ _ blowing in the _ _ _ _.  


First off...

More stuff about passwords? Ugh. Beat a dead horse much?

But I cannot, in good faith, call myself a “usable security” researcher without doing a paper on passwords. So, here it is.
Sometimes, we need strong secrets

Password managers

Encrypted drives

Journalists / whistle blowers
Strong passwords

If we still need strong passwords, we should make strong passwords easy to use.

Typical approach: Teach people strong password generation strategies. But people aren’t random enough.

Out-of-vogue approach: Give people randomly assigned strong passwords. But, people can’t remember anything...right?
Actually, it works...

Bonneau & Schechter (2014) conditioned lay people to learn strong assigned secrets (56.4 bits) incrementally.

On login, had people enter assigned secret (which is shown after a delay that gradually increases).

People could avoid delay if they enter the assigned secret from memory.

Example: apple beetle crane dog ear flame
User ID: 
hotshot@handsome.cc

Password: 
apple beet
User ID:
hotshot@handsome.cc

Password:
apple beetle crane dog ear flame
apple beet
94% of participants memorized all of their secret words.
But...

The process was long (required 36 sessions)

Recall fell sharply after 2 weeks (62% remembered)
Can we do better?

Ideally, the process should:

- Require fewer training sessions
- Retained for longer
Mnemonics?!

“Strategies to enhance the learning and recall of information”

Most useful when memorizing lists.

Basic idea: Transform abstract, difficult to visualize concepts into tangible, easy to visualize concepts.
Memorize: apple, beetle, crane

<table>
<thead>
<tr>
<th>Chain-type</th>
<th>Peg-type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weave list items into narrative chain.</td>
<td>Peg list items onto cueing structure.</td>
</tr>
<tr>
<td>an apple being eaten by a beetle while being lifted up by a crane.</td>
<td>one is nun a nun throwing an apple</td>
</tr>
<tr>
<td></td>
<td>two is shoe a shoe squashing a beetle</td>
</tr>
<tr>
<td></td>
<td>three is tree a tree being lifted by a crane</td>
</tr>
</tbody>
</table>
Research Question:

Can we use mnemonics to make learning strong secrets quicker and longer-lasting?
Our contribution

1. Two computer-assisted mnemonic training schemes designed to incrementally teach people strong secrets.

2. Experimental evaluation of these training schemes relative to the rote baseline akin to prior work.
Method
Method Overview

Recruit participants, via MTurk, for a recurring “attention” test over 10 days.

Assign participants, randomly, to a treatment or control group.

Teach participants a randomly selected strong secret with an assigned training regimen.

Compare, across training regimens, how fast participants learn their secrets and how well they remember their secret words 2-weeks after the fact.
Recruitment

Participants solicited via MTurk to take “attention test”.

“Attention test” cover story used to distract participants from focusing on learning their secret words.

Participants had to return to do 45 attention tests over maximum of 10 days with at least 1 hour between each.
Assignment

Participants were randomly assigned to one of five experimental groups:

**Story:** Mnemonic treatment.

**Peg-Word:** Mnemonic treatment.

**Feedback:** Active learning treatment.

**Rote:** Control. Baseline from prior work.

**Dropout:** Control. No secret words.
Assignment

Participants were randomly assigned to one of five experimental groups:

**Story**: Mnemonic treatment.

**Peg-Word**: Mnemonic treatment.

**Rote**: Control. Baseline from prior work.

**Feedback**: Active learning treatment.

**Dropout**: Control. No secret words.
Teach secret

**Secret Encoding:** 6 nouns randomly selected out of a broader set of 676 options (56.4 bits).

Assigned in two, 3-word chunks.

Presented one chunk at a time — only shown second chunk after “learning first” (entering 3x consecutively from memory)
Teach secret

apple  beetle  crane

apple  beetle  cra

On each login, participants had to enter their secret words.

If they couldn’t remember, they would eventually be shown their words to copy.
Comparison

Primarily interested in two comparisons of mnemonic vs rote baseline:

- **Number of learning sessions required**: Lower is better (previous study rote method: median of 36)

- **2-week recall rate**: Higher is better (previous study rote method: 62%)
Experimental Groups
3 Groups

**Story Mnemonic:** Users wrote two sentences, each containing three of their secret words, in order.

**Peg-Word Mnemonic:** Users were given public “peg” words and created sentences linking each peg to a secret.

**Rote Baseline:** No special training. Users just had to manually enter words to the best of their memory.

*Dimensions of Variance:* mnemonic creation, chunk rehearsal, and hint progression.
Rote Baseline

**Mnemonic Creation:** none

**Chunk Rehearsal:** Text-boxes with hints ultimately rendered on top for copying.

**Hint Progression:** Started at 0 seconds, progresses up to a maximum of 10 seconds with a 0.5 second increase every subsequent attempt.

apple  beetle  crane
apple  beetle  cra
Story Mnemonic

Mnemonic Creation: Asked participants to write a visual, memorable sentence stringing together all three words.

Example: cat, leaf, wind => A pink cat is chasing a giant leaf blowing in the wind.

Participants had to comply to proceed — our app checked whether the story sentence was valid.
Story Mnemonic

Chunk Rehearsal: Three-phases of chunk rehearsal to gradually wean participants off mnemonic assistance.

Introduced the *hint well* to facilitate this gradual transition. The hint well allowed participants to enter their secret words *inline*.

**Phase 1: Full assistance**

A pink _ _ _ _ is chasing a giant _ _ _ _ blowing in the _ _ _ _. 
Story Mnemonic

Phase 2: Reduced assistance

Phase 3: No assistance
Story Mnemonic

*Phase Transition*: Start in full-assistance, then reduced-assistance, then no-assistance. Transition after first entering from memory in previous phase.

*Hint Progression*: Start at 1 second, max of 10 seconds. Increment by 0.5 seconds on each attempt.
Peg-Word Mnemonic

**Mnemonic Creation:** First, asked participants to select rhyme-based peg words. Then, asked participants to write three visual, memorable sentences stringing together public peg words with secret words.

**Example:** cat, leaf, wind

1 is \{nun, bun, gun, sun\}

2 is \{shoe, shrew, zoo, screw\}

3 is \{tree, bee, key, sea\}
Peg-Word Mnemonic

<table>
<thead>
<tr>
<th>Public peg word</th>
<th>Secret word</th>
<th>Mnemonic hint sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>nun</td>
<td>cat</td>
<td>The nun pulls the screeching cat off of her head.</td>
</tr>
<tr>
<td>leaf</td>
<td>shoe</td>
<td>The leaf lands gracefully in the worn old shoe.</td>
</tr>
<tr>
<td>wind</td>
<td>tree</td>
<td>The wind has no effect on the steadfast tree.</td>
</tr>
</tbody>
</table>

Participants had to comply to proceed — our app checked whether the peg sentences were valid.
Peg-Word Mnemonic

Chunk Rehearsal: Same three-phases of chunk rehearsal to gradually wean participants off mnemonic assistance. *Difference*: hint sentences were shown per word instead of per chunk.

**One is nun.** The nun pulls the screeching _ _ _ _ off of her head.

| cat | leaf | wind |

But hints still operated per-chunk, not per-word.
Peg-Word Mnemonic

Phase Transition: Same as story.

Hint Progression: Same as story.
Results

Learning Speed & Memorability
Some Descriptive Stats

351 participants signed up.

242 participants finished.

31 years old on average (sd 10).

52 percent female.

<table>
<thead>
<tr>
<th></th>
<th>Rote</th>
<th>Story</th>
<th>Peg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assigned</td>
<td>71</td>
<td>81</td>
<td>75</td>
</tr>
<tr>
<td>Completed</td>
<td>51</td>
<td>57</td>
<td>48</td>
</tr>
<tr>
<td>2-week return</td>
<td>42</td>
<td>48</td>
<td>43</td>
</tr>
</tbody>
</table>
Learning sessions

Participants who entered their secret words, from memory, three consecutive times had *learned* their secret words.

The number of sessions *prior* to that 3-chain were “learning sessions”.

**Q1:** How did the number of learning sessions differ between those in our mnemonic treatments versus those in the rote baseline?
Both story (7.5) and peg-word (9) participants required significantly fewer learning sessions than rote (12).
2-Week Recall Rate

Q2: How did the 2-week recall rate vary between participants in the mnemonic treatments versus the rote baseline?
# 2-Week Recall Rate

<table>
<thead>
<tr>
<th>Name</th>
<th>Entropy</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perfect</td>
<td>56.4</td>
<td>apple beetle crane dog ear flame</td>
</tr>
<tr>
<td>Single Swap</td>
<td>53.8</td>
<td>apple crane beetle dog ear flame</td>
</tr>
<tr>
<td>Forgot One</td>
<td>49.6</td>
<td>apple ? crane dog ear flame</td>
</tr>
<tr>
<td>Relaxed Order</td>
<td>46.9</td>
<td>ear apple dog beetle flame crane</td>
</tr>
</tbody>
</table>
2-Week Recall Rate

<table>
<thead>
<tr>
<th></th>
<th>Story</th>
<th>Peg-Word</th>
<th>Rote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perfect</td>
<td>75%</td>
<td>42%</td>
<td>60%</td>
</tr>
<tr>
<td>Single Swap</td>
<td>84%</td>
<td>47%</td>
<td>65%</td>
</tr>
<tr>
<td>Forgot One</td>
<td>84%</td>
<td>67%</td>
<td>74%</td>
</tr>
<tr>
<td>Relaxed Order</td>
<td>86%</td>
<td>47%</td>
<td>65%</td>
</tr>
</tbody>
</table>

Takeaways

**Story** did significantly better than Rote.

**Peg-Word,** surprisingly, did worse (but non-significant).

\[ p < 0.05 \]
\[ p < 0.10 \]
Results Summary

**Mnemonics** require fewer learning sessions and can provide recall benefits, but only if not too complicated.

**Story** mnemonic did best, providing significant improvements over rote on both reducing required learning sessions and 2-week recall.

We can perhaps improve retention further if we also facilitate remembering the mnemonics themselves.
Conclusion
Some use-cases warrant the use of provably strong secrets.

Bonneau and Schechter (2014): Lay people can reliably learn such secrets through conditioning, but it takes a while and recall falls sharply after extended disuse.

We designed and implemented two computer-assisted mnemonic training regimens to speed-up learning and increase recall rates after extended disuse.

Our **story** mnemonic did both of those things, but our **peg-word** mnemonic was surprisingly ineffective.
Testing Computer-Aided Mnemonics and Feedback for Fast Memorization of High-Value Secrets

Sauvik Das  
Carnegie Mellon University

Jason Hong  
Carnegie Mellon University

Stuart Schechter  
Microsoft Research

**Key-takeaway:**
Our computer-assisted story mnemonic decreased the required number of sessions required to learn and increased the two-week recall rates of strong, randomly assigned secrets.

Email contact: sauvik@cmu.edu