Testing the KRK theory

Frank E. Ritter, Jake Oury, Farnaz Tehranchi

acs.ist.psu.edu/papers
frank.ritter@psu.edu
21 july 2018

Slides presented at the ACT-R Workshop, July 2018

WIRES ACT-R review

Thanks to: Ashley McDermott, Sarah Stager, Peter Weyhrauch, Ben Bauchwitz, N00014-15-1-2275
Project Objectives

Improve learning and retention in training by understanding learning better:

• Develop a cognitive architecture-based theory to optimize training and retention
• Explore and design better training schedules
• Create deeper understanding of maintenance skill learning and retention
• Implement tutoring approach to apply results

Technical Approach

• A learning and retention theory
• Learn more about how learning schedules interact with tasks
• System to build adaptive tutors quickly (D2P2) using learning theory, HCI methods, software engineering, + example tutors
• Refine D2P2 tutoring system iteratively through use

Accomplishments/Impact/Transitions

1. Microgenetic analysis of learning & retention
2. Work on training schedules study finished
3. D2P tutoring architecture
   * Web deployed w/ simulations
   * Supports mobile devices
   * Adaptive instruction with page annotations
   * Usable by undergraduates
   * Provided to NSMRL for feedback
4. Built tutors (Revised: CLS, MTT),
   (New: Maintenance+Mends simulation, Medals, Rate&Ratings, Tutor², Chess pieces)
   (Other: AF Trauma Nursing, DHA AAJT)
5. D2P approach used in Canadian DRDC study
Technical Approach 1: KRK theory

(Kim, Ritter, & Koubek, 2013)
TA 2: Piloting study with ACT-R model

Dismal spreadsheet task (14 subtasks, 4 repetitions, 1 delayed test)

Nice model to data (N=30) fit

Mean

Declarative Memory HTA

(📖 Dismal, a spreadsheet in Emacs
Ritter & Wood, 2005)

(📖 RUI, a keystroke logger,
Kukreja et al. 2006; Morgan et al. 2013)

(📖 Keystroke vs. GUI,
Kim & Ritter, 2015)

(📖 Herbal model test,
Paik, Kim, Ritter, & Reitter, 2015)
TA2: Piloting study with ACT-R

Figure 3: Predictions for task time with OL (black). Forgetting curves show task time after [1-5] days of decay after a period of consistent practice.

Figure 5: Predictions for task time with BL for ten trials of practice (black). Forgetting curves show task time after [1-5] days of decay after a period of consistent practice.

Thus (?)

• 1, 2, 4-5 trials w/ Delay of 3, 5, 7 days
• No relearning from test
• 24 h delay makes a difference
• Problems with decay

Figure 8. Predictions for task completion time with BL with 24-hour decay periods between each trial and an initial 24 decay on the declarative knowledge, adjusted

(© Oury, Tehranchi, & Ritter, 2018)
TA2: Piloting study eyes and hands

• Revision of SegMan from C to Java, using Robot and Sikuli libraries
• Model does the task with full, uninstrumented interaction
• Does the Dismal task, like, it *actually* does it
• Found mistakes in model b/c we could see
• Closer fit:

<table>
<thead>
<tr>
<th>Day</th>
<th>Human</th>
<th>Original Model</th>
<th>JSegMan Hands+Eyes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1366 (60.8)</td>
<td>1326 (12.1)</td>
<td>1339 (11.7)</td>
</tr>
<tr>
<td>2</td>
<td>894 (26.6)</td>
<td>891 (6.1)</td>
<td>894 (6.5)</td>
</tr>
<tr>
<td>3</td>
<td>727 (25.5)</td>
<td>693 (4.5)</td>
<td>704 (5.0)</td>
</tr>
<tr>
<td>4</td>
<td>659 (22.7)</td>
<td>594 (5.8)</td>
<td>614 (4.4)</td>
</tr>
<tr>
<td>R2</td>
<td>.997</td>
<td>.9984</td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>1745</td>
<td><strong>820</strong></td>
<td></td>
</tr>
</tbody>
</table>

(Tehranchi & Ritter, 2018a, b)
TA2: Study proposed

1. Subjects, N=135 PSU students balanced by major

2. Tasks:
   a) Complex troubleshooting task (45 min.)
      (BenFranklinRadar = Diag++++), in D2P tutor
      5x bigger than DiagTask 35v.7
      45 faults, single, multiple
   b) KLM constants tasks (4 min. at start), vertical mouse
   c) At test Knowledge types tasks (3x4min.)
      i. Tasks related to the complex task, procedural troubleshooting
      ii. P/M (mouse/keystroke speed)
      iii. Recognition (10 stimuli, 10 foils, have you seen this before?)
   d) Record task actions and mouse & keystrokes

3. Design:
   3 training amounts (1, 2, 4) before retention measure
   3 retention amounts (3, 5, 7 days), without feedback
   x 15 Ss = 135 Ss, 450 sessions

4. Data will provide trial#s and times for entry into each stage for a task or several tasks based on decay rates
Device Schematic

MENDS v.3.6 (7/3/2018)

Antenna System:
- Antenna
- Transmitter
- Receiver
- Receiver transmission
- Transmitter

RF amplifier 1
- Stage 1
- Frequency

RF amplifier 2
- Stage 2

RF amplifier 3
- Stage 3

Power Supply
- Limiter
- Overdriven
- Sin. wave oscillator
- Adjuster

Synchronizer

Receiver:
- AFC
- AFC frequency
- AFC mixer
- Local
- Control
- IF amplifier
- Video

Sweep control
- Sweep generator
- Gate
- Intensity gate

Indicator

Mechanical bearing information

* = indicator light for cabinet
TA2: Study proposed

1. Subjects, N=135 PSU students balanced by major

2. Tasks:
   a) Complex troubleshooting task (45 min.)
      (BenFranklinRadar = Diag++++), in D2P tutor
      5x bigger than DiagTask 35v.7
      45 faults, single, multiple
   b) KLM constants tasks (4 min. at start), vertical mouse
   c) At test Knowledge types tasks (3x4min.)
      i. Tasks related to the complex task, procedural troubleshooting
      ii. P/M (mouse/keystroke speed)
      iii. Recognition (10 stimuli, 10 foils, have you seen this before?)
   d) Record task actions and mouse & keystrokes

3. Design:
   3 training amounts (1, 2, 4) before retention measure
   3 retention amounts (3, 5, 7 days), without feedback
   x 15 Ss = 135 Ss, 450 sessions

4. Data will provide trial#s and times for entry into each stage for a task or several tasks based on decay rates
Summary/Questions

• Model used to design a study
• Model can interact via OS
• Will generate single set of 4 learning and retention curves for a complex task
• Corrections, citations, comments, requested
REFERENCES

(Materials not in list of resulting publications)


Paik, J., & Ritter, F. E. (2016). Evaluating a range of learning schedules: Hybrid training schedules may be as good or better than distributed practice for some tasks. Ergonomics. 59(2). 276-290.


