Testing the KRK theory

Frank E. Ritter, Jake Oury, Farnaz Tehranchi

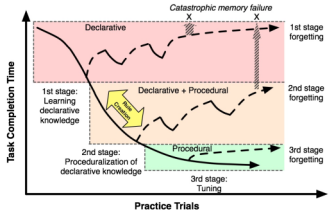
acs.ist.psu.edu/papers

frank.ritter@psu.edu

21 july 2018

Slides presented at the ACT-R Workshop, July 2018

WIRES ACT-R review



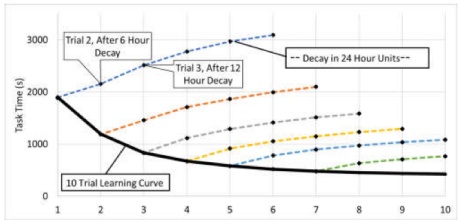
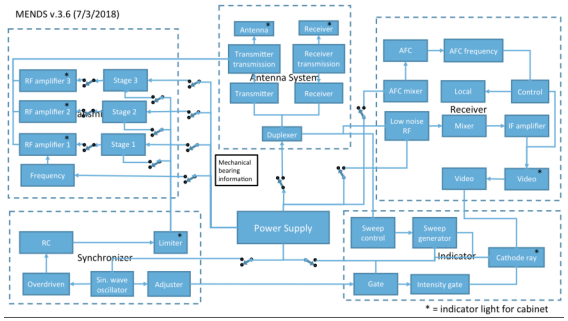


Figure 8. Predictions for task completion time with BL with 24-hour decay periods between each trial and an initial 24 decay on the declarative knowledge, adjusted



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

EXPANDING SMART TUTORING TO SUPPORT SKILL LEARNING AND RETENTION

Frank Ritter (Penn State), Peter Weyhrauch (Charles River Analytics)

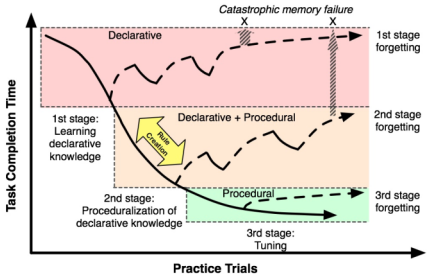
1 Apr 15 – 30 Mar 2020, started 28 Jul 2015

Project Objectives

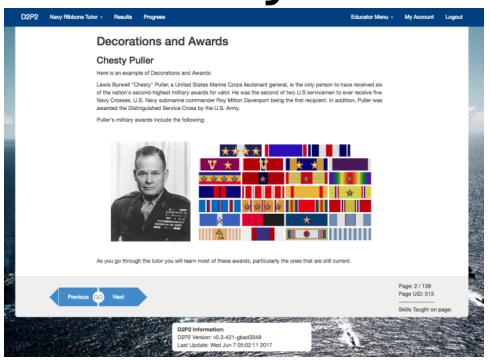
Improve learning and retention in training by understanding learning better:

- 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

KRK Theory



D2P2 System



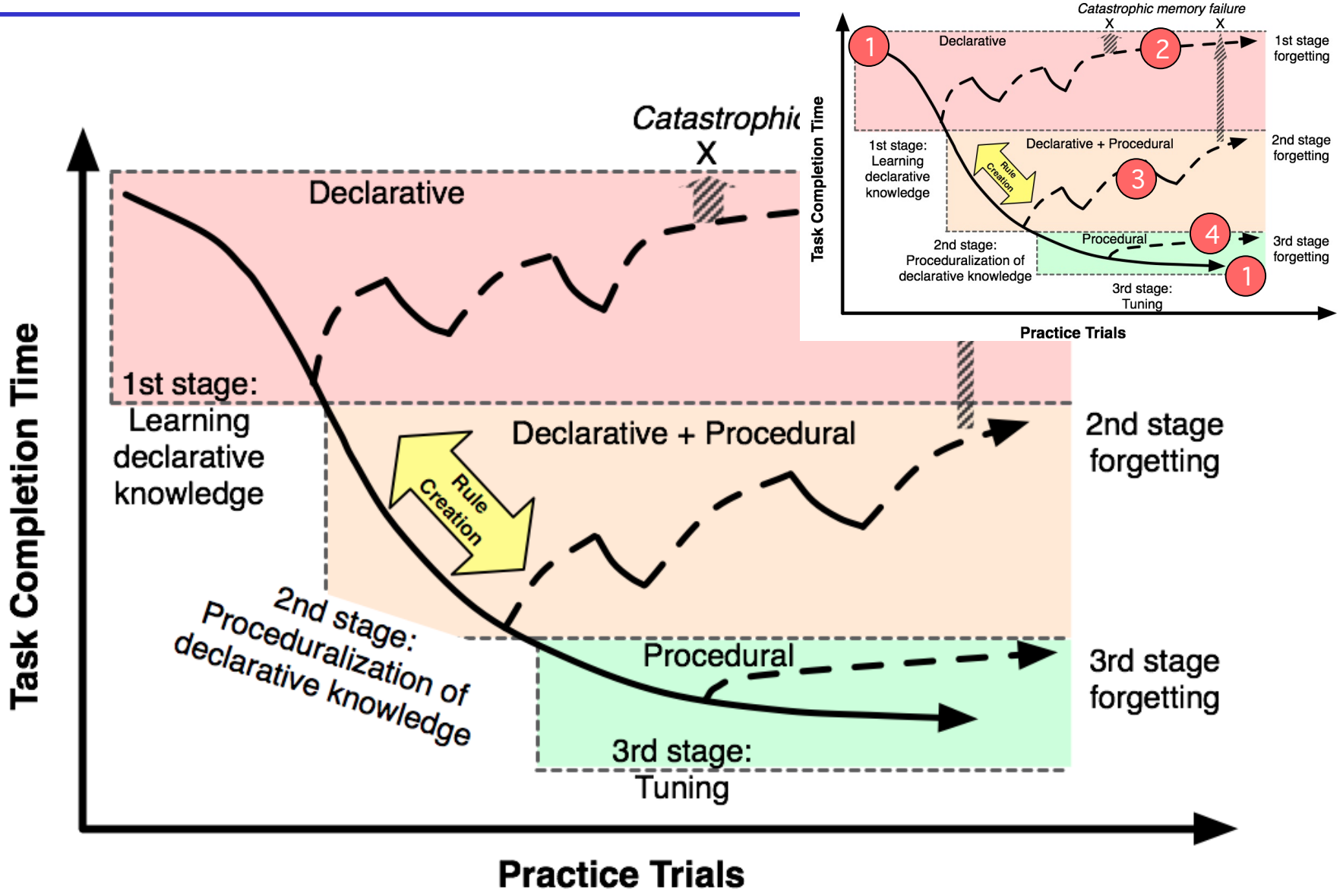
Technical Approach

- 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, Tutor^2, Chess pieces)
   (Other: AF Trauma Nursing, DHA AAJT)
5. D2P approach used in Canadian DRDC study

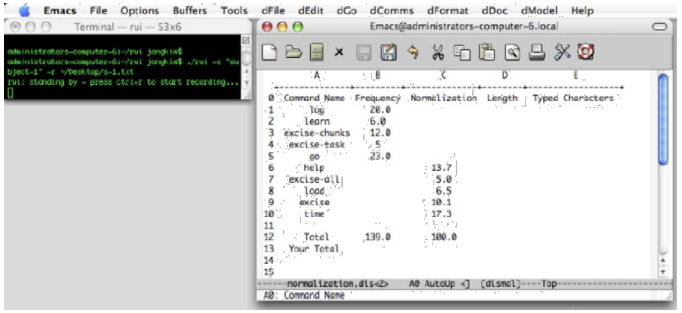
Technical Approach 1: KRK theory

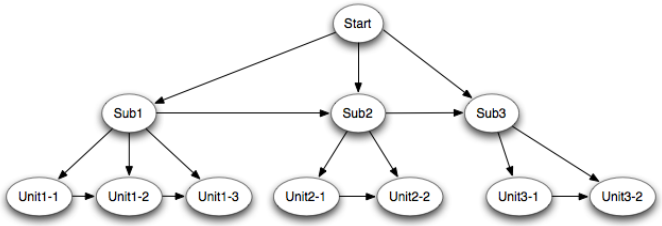


📖(Kim, Ritter, & Koubek, 2013)

TA 2: Piloting study with ACT-R model

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



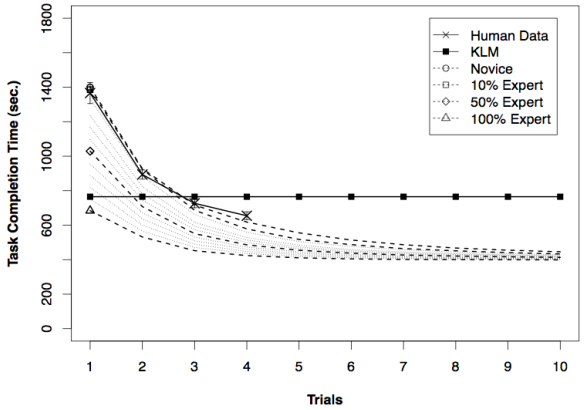


Declarative Memory HTA

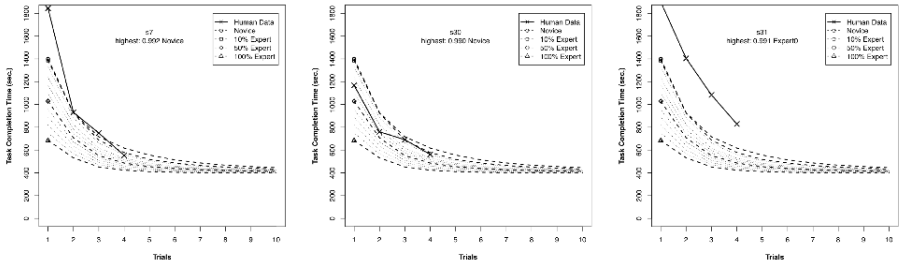
(📖 Dismal, a spreadsheet in Emacs
Ritter & Wood, 2005)
(📖 RUI, a keystroke logger,
Kukreja et al. 2006;
Morgan et al. 2013)

Nice model to data (N=30) fit

Mean



Avg



(📖 Keystroke vs. GUI,
Kim & Ritter, 2015)
(📖 Herbal model test,
Paik, Kim, Ritter, & Reitter, 2015)

TA2: Piloting study with ACT-R

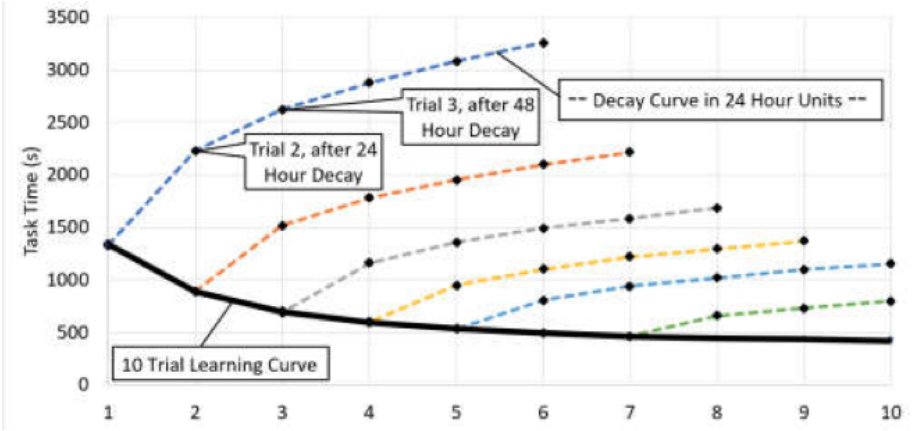


Figure 3: Predictions for task time with OL (black). Forgetting curves show task time after [1-5] days of decay after a period of consistent practice.

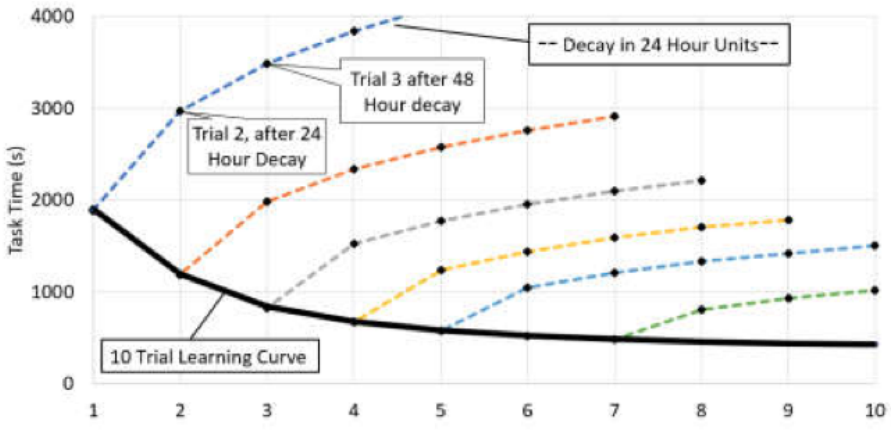


Figure 5: Predictions for task time with BL for ten trials of practice (black). Forgetting curves show task time after [1-5] days of decay after a period of consistent practice.

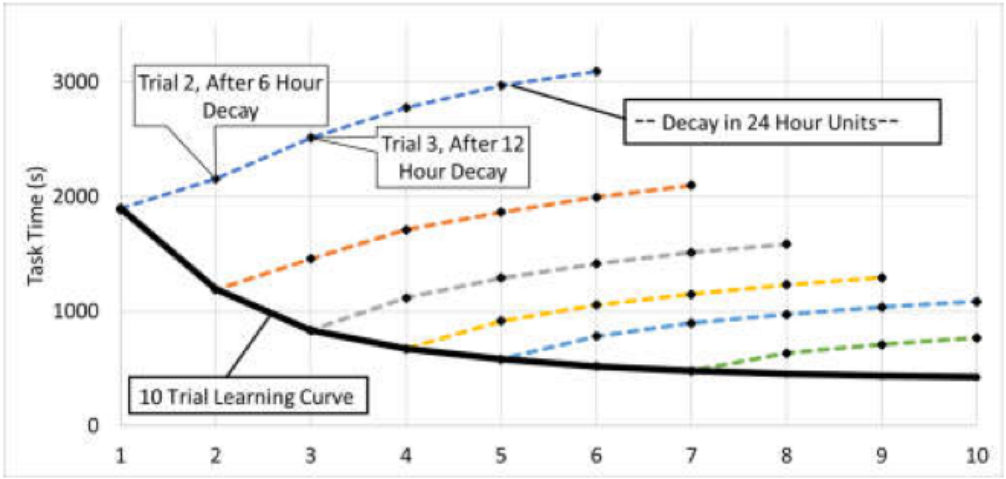


Figure 8. Predictions for task completion time with BL with 24-hour decay periods between each trial and an initial 24 decay on the declarative knowledge, adjusted

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

(📖 Oury, Tehranchi, & Ritter, 2018)

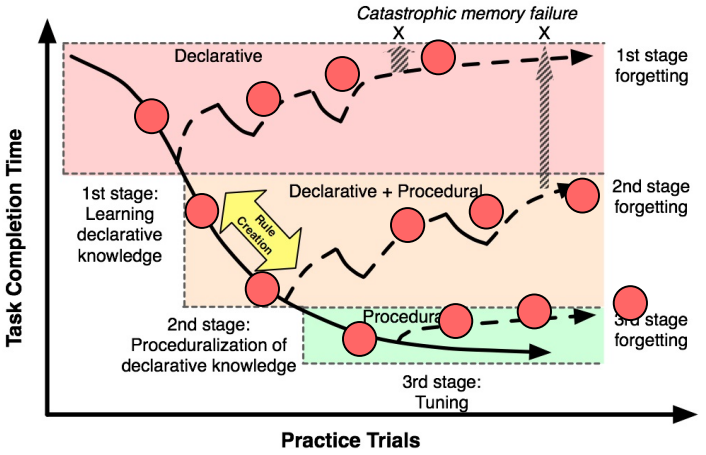
TA2: Piloting study eyes and hands

- 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:

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

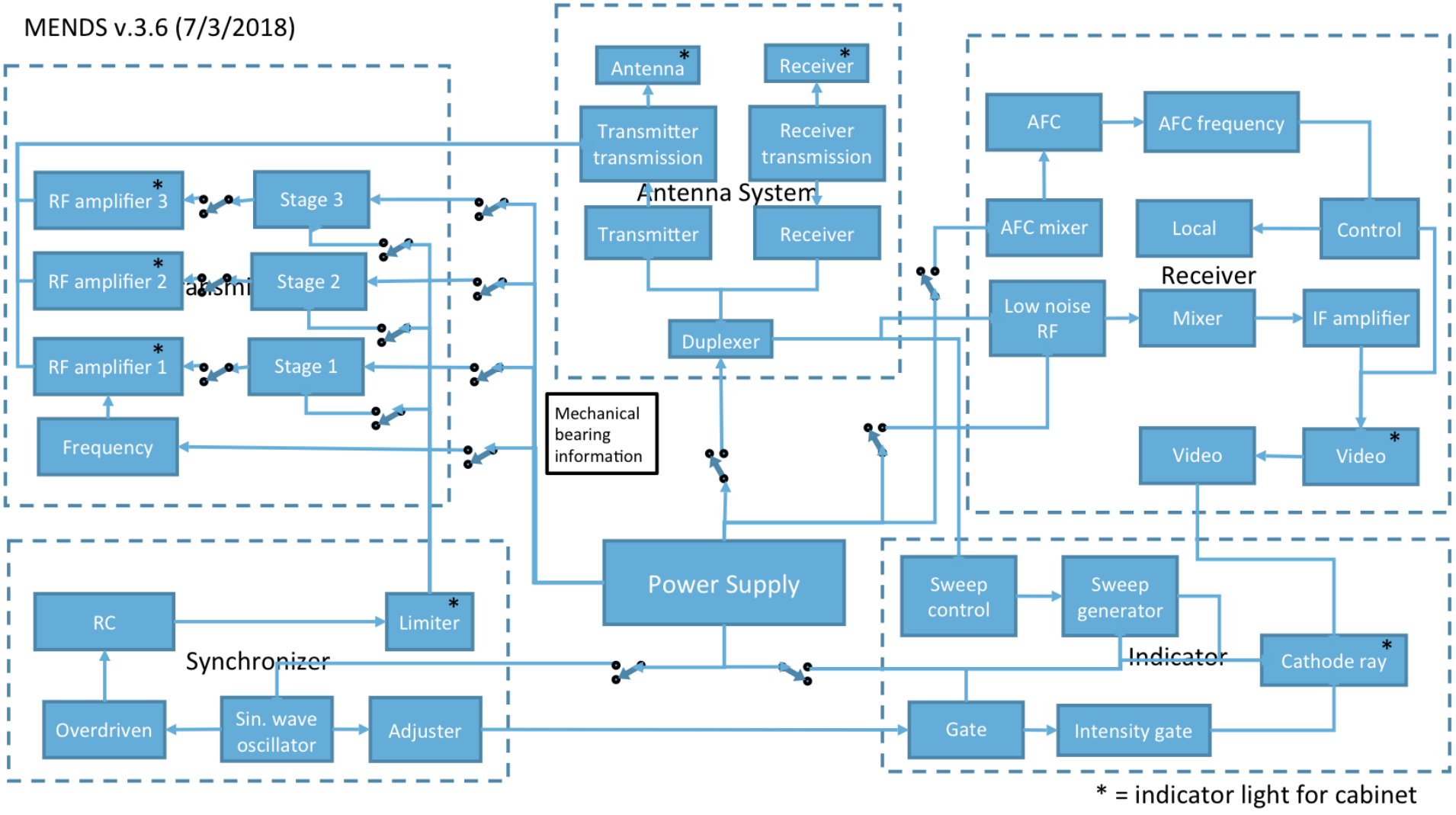
(📖 Tehranchi & Ritter, 2018a, b)

TA2: Study proposed

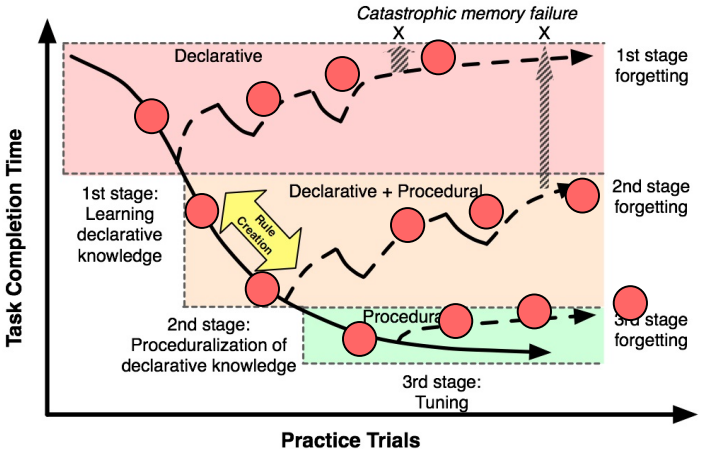


1. Subjects, N=135 PSU students balanced by major
2. Tasks:
   a) Complex troubleshooting task (45 min.)
      (BenFranklinRadar = Diag++++), in D2P tutor
      5x bigger than DiagTask 35v.7
      45 faults, single, multiple
   b) KLM constants tasks (4 min. at start), vertical mouse
   c) At test Knowledge 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



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

(📖 acs.ist.psu.edu/papers/)

(Materials not in list of resulting publications)

📖 Kim, J. W., & Ritter, F. E. (2015). Learning, forgetting, and relearning for keystroke- and mouse-driven tasks: Relearning is important. Human-Computer Interaction, 30(1), 1-33.

📖 Kim, J. W., Ritter, F. E., & Koubek, R. J. (2013). An integrated theory for improved skill acquisition and retention in the three stages of learning. Theoretical Issues in Ergonomics Science, 14(1), 22-37.

📖 Kukreja, U., Stevenson, W. E., & Ritter, F. E. (2006). RUI-Recording User Input from interfaces under Windows and Mac OS X. Behavior Research Methods, 38(4), 656-659.

📖 Morgan, J. H., Cheng, C.-Y., Pike, C., & Ritter, F. E. (2013). A design, tests, and considerations for improving keystroke and mouse loggers. Interacting with Computers, 25(3), 242-258.

📖 Oury, J. D., Tehranchi, F., & Ritter, F. E. (2018). Predicting learning and retention of a complex task. ICCM 2018.

📖 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.

📖 Paik, J., Kim, J. W., Ritter, F. E., & Reitter, D. (2015). Predicting user performance and learning in human-computer interaction with the Herbal compiler. *ACM Transactions on Computer-Human Interaction, 22*(5), Article No.: 25.

📖 Ritter, F. E., & Wood, A. B. (2005). Dismal: A spreadsheet for sequential data analysis and HCI experimentation. *Behavior Research Methods, 37*(1), 71-81.

📖 Tehranchi, F., Ritter, F. E. (2018a). Using Java to provide cognitive models with a more universal way to interact with graphical user interfaces. The Proceedings of the International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation.

📖 Tehranchi, F., Ritter, F. E. (2018b). Modeling visual search in interactive graphic interfaces: Adding visual pattern matching algorithms to ACT-R. To appear in the Proceedings of ICCM 2018.