Predicting learning and retention in a complex task

David Peebles
Department of Psychology, University of Huddersfield, UK.

Centre for Cognition and Neuroscience

Competence retention and training

- To maintain high levels of task knowledge and skill performance, military personnel undergo regular testing and refresher training to combat effects of *skill fade*
- Retraining schedules are often standardised (e.g., calendar-based) **but** individual differences in learning and retention can result in people performing below threshold long before scheduled retraining session
- Evidence that skill retention varies according to the cognitive components involved (e.g., Cahillane et al., 2013). Refresher training may be optimised by:
  - Analysing the knowledge, skills and attitudes involved in the task
  - Measuring individual differences in learning and retention

Competence retention analysis

Cahillane et al., (2013) propose a three-level categorisation of retention, defined on a criterion value of 50% competence after a given period of time since the last training session:

High > 50% competent after 12 months
Moderate 50% competent after 5 months
Low 50% competent after 2 months

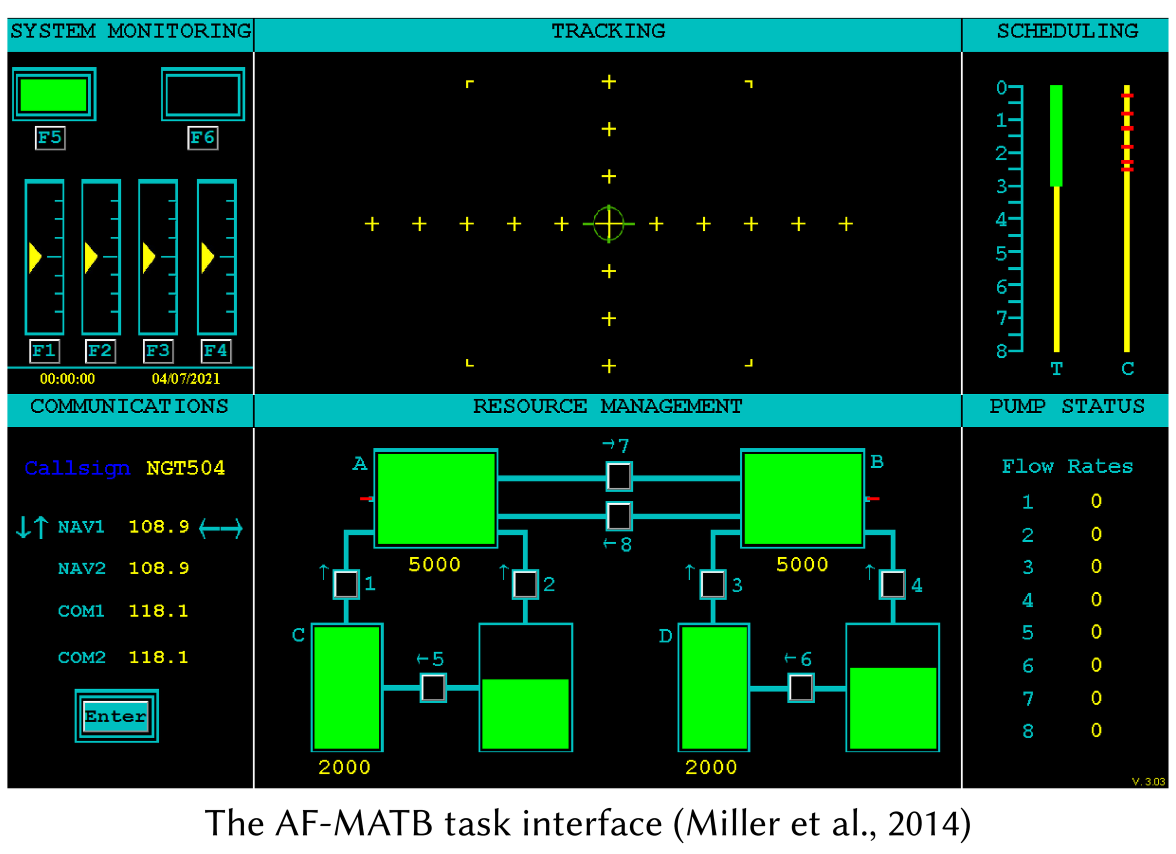
Psychological Domain	Description	Application Frequency	Retention Level	MATB Sub-task
Continuous Psychomotor skills	Perceptual motor tasks such as target tracking, driving a vehicle and manipulating a weapon.	Very Frequent	High	Tracking
		Moderately Frequent	High	
Explicit Knowledge	Memory for facts, concepts, information and theories.	Infrequent	High	
Discrete psychomotor skills	A combination of physical and procedural skill e.g., weapon handling.	Very Frequent	High	
		Moderately Frequent	Moderate	
Decision-making skills	Reasoning in order to identify a problem and correct course of action.	Infrequent	Moderate	Resource Management
Procedural skills	Steps within drill/sequence and order sequence of tasks.	Very Frequent	Moderate	
		Moderately Frequent	Low	Communication
		Infrequent	Low	

The Predictive Performance Equation (PPE)

- A set of mathematical equations that describe the effect of three factors that affect knowledge acquisition and retention (and subsequent task performance):
  - Amount of practice (frequency effect)
  - Amount of time since last practice session (recency effect)
  - The temporal distribution of practice (spacing effect)
- The PPE predicts an individual's future performance based on the previous training history (Jastrzembski & Gluck, 2009; Walsh et al., 2018). Method:
  - Gather time stamped performance data from individuals during training
  - Calibrate PPE by finding parameters values that maximise correspondence between model output and individual performance data
  - Use the calibrated model to predict each individual's future performance and make training prescriptions

NASA Multi-Attribute Task Battery (MATB)

- Complex task with four sub-tasks
- Version used created by AS Air Force (AF-MATB)
- Used extensively in human factors research into multitasking



The AF-MATB task interface (Miller et al., 2014)

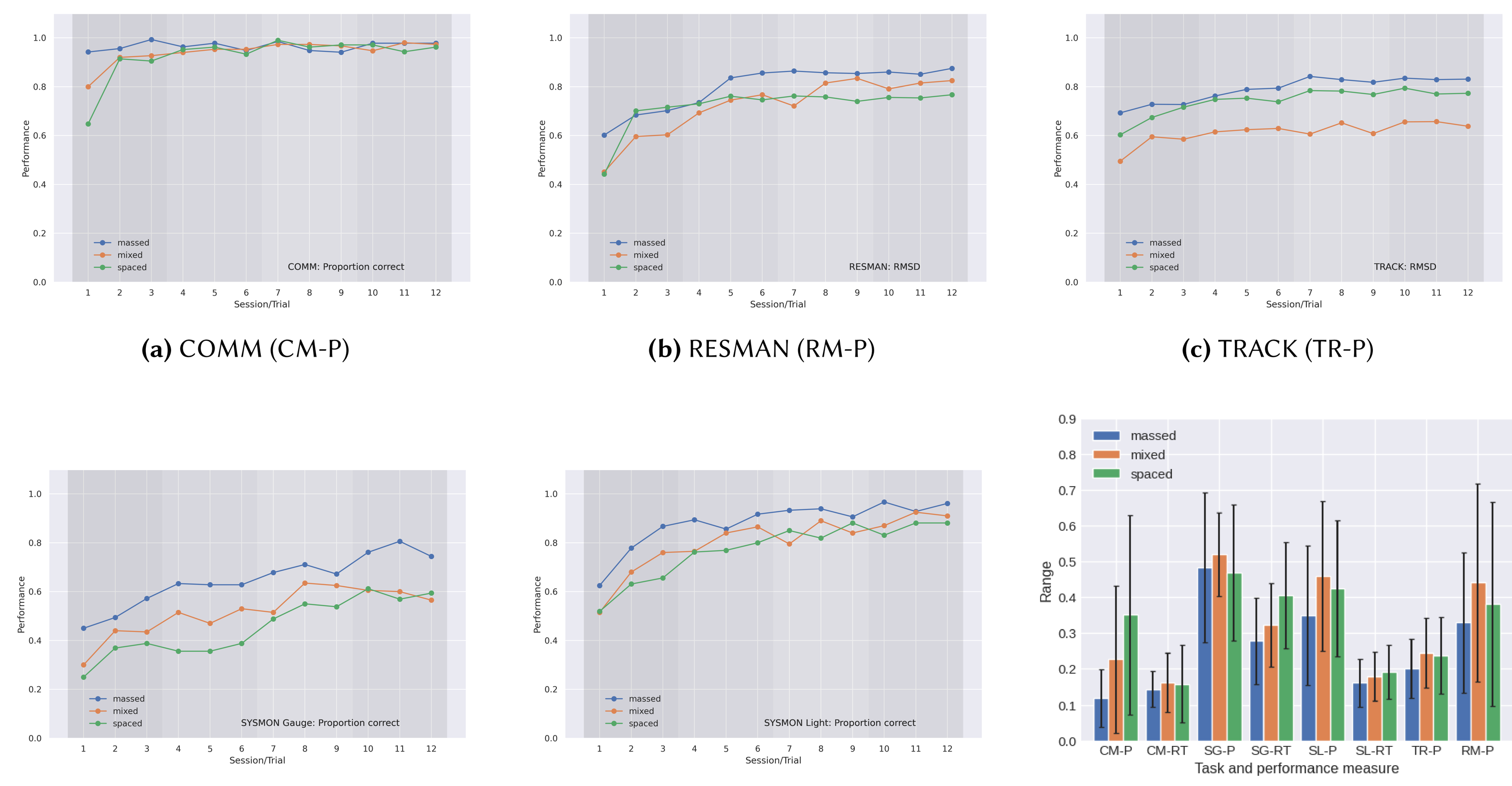
System monitoring (SYSMON) Proportion correct. Mean RT
Tracking (TRACK) RMSD, crosshair and target distance
Communication (COMM) Proportion correct responses. Mean RT
Resource management (RESMAN) RMSD, actual and target fuel levels

Experiment

- Participants: 27 staff and students from University of Huddersfield
- 4 training sessions over 4 weeks, 3 different training schedules:
  - **Spaced:** 1 day per week for 4 weeks (8 participants)
  - **Massed:** 4 consecutive days in 1 week (9 participants)
  - **Mixed:** 2 days for 2 weeks, 1 week separation (10 participants)
- Each session: Three 10-minute trials separated by 5-minute breaks. 12 data points in total (three trials × four training sessions)
- 2 test sessions: 6 and 12 weeks after last training session

	Weeks															
Condition	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Massed				4						1						1
Spaced	1	1	1	1						1						1
Mixed		2		2						1						1

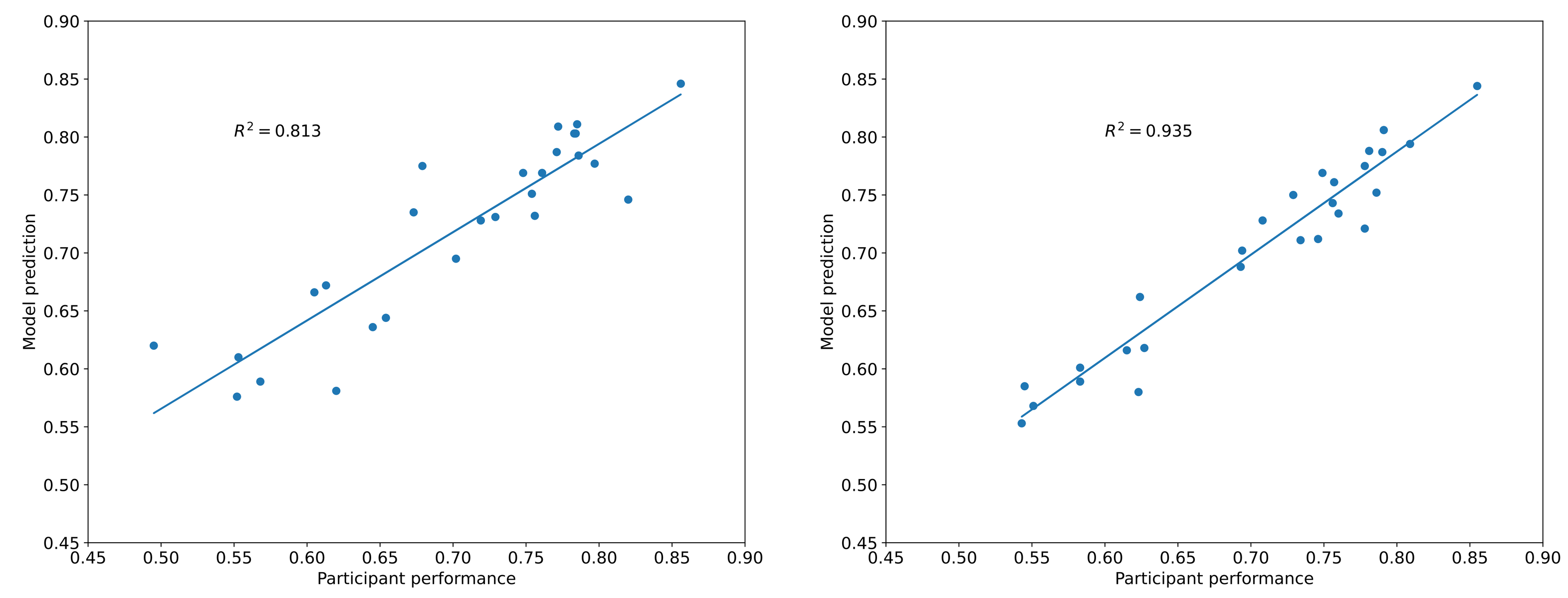
Human performance on the MATB subtasks



(a) COMM (CM-P) (b) RESMAN (RM-P) (c) TRACK (TR-P)
(d) SYSMON Gauge (SG-P) (e) SYSMON Light (SL-P) (f) Performance ranges

Model predictions for the two test sessions

- The subtask measures were transformed onto a common scale and then averaged for each participant to create a single, global MATB score
- The PPE's predictions were tested on sessions 5 and 6, (approximately 43 and 86 days respectively after the fourth training session)
- For each test session, the model was fitted to the individual's performance data from the previous sessions and thhe fitted model was then used to predict performance at the date and time of the first trial of the test session



Session 5 (test session 1) Session 6 (test session 2)

Conclusions

- Experiment has generated a rich dataset of individual learning and forgetting in a complex task involving multiple sub-tasks
- Experiment provides additional support for the PPE. The model produces accurate predictions over retention intervals ranging from 27 to 111 days
- Learning differences not consistent with CRA classification
- Additional analysis from sessions 5 and 6, combined with a detailed task analysis, may provide further insight into differences in retention over longer intervals
- Further details and data available at OSF: https://osf.io/uc4fy

References

- Cahillane, M. A., Launchbury, C., MacLean, P., & Webb, S. (2013). *Competence retention* (tech. rep. DHC-STC_12_T_T2_001_1.1/005 V5.0). Defence Science and Technology Laboratory.
- Jastrzembski, T. S., & Gluck, K. A. (2009). A formal comparison of model variants for performance prediction. In A. Howes, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 9th International Conference on Cognitive Modeling*.
- Miller, W. D., Schmidt, K. D., Estepp, J. R., Bowers, M., & Davis, I. (2014). *An updated version of the US Air Force multi-attribute task battery (AF-MATB)* (tech. rep. AFRL-RH-WP-SR-2014-0001). Air Force Research Laboratory.
- Walsh, M. M., Gluck, K. A., Gunzelmann, G., Jastrzembski, T. S., & Krusmark, M. (2018). Evaluating the theoretic adequacy and applied potential of computational models of the spacing effect. *Cognitive Science*, 42, 644-–691.