Taxonomy and IMPRINT modeling

Bill Raymond

University of Colorado, Boulder

MURI Annual Review
August 14, 2009
Overview

- **Taxonomy**
  - Performance shaping functions for IMPRINT
  - Quantitative framework for training

- **IMPRINT modeling**
  - RADAR task modeling of Experiment 1 completion
Performance shaping functions

- **Last year**
  
  - Listed principles with empirical data
  - Used RADAR experiments as example of source data
  - Focused on quantification procedure for 4 principles
  - Showed functions for these principles, parameters unspecified
  - Listed IMPRINT taxons affected for each

- **This year**
  
  - Focus on 2 principles from different classes
  - Quantify principles to provide performance shaping functions
  - Derive parameter values from experimental data
  - Provided to Charneta Samms of ARL
Task difficulty

Desirable difficulty during declarative learning...

(a) lowers accuracy during learning
(b) improves long-term retention

\[
\begin{align*}
    p_1 &= (1+a) \ p_L \\
    p_2 &= (1+b) \ p_{DP}
\end{align*}
\]

From Schneider, Healy, and Bourne (2002) – foreign vocabulary learning

\[a = -0.37; \ b = 0.23\]

IMPRINT task taxons affected: information processing numerical analysis
Mnemonic procedures: Use of prior knowledge

Relating facts to be learned to already well-known facts...

(a) improves performance during learning \( p_1 = (1+c) \ p_L \)
(b) improves performance after a delay. \( p_2 = (1+d) \ p_{DP} \)

From Kole and Healy (2009) – “friends and families” experiment
\( c = 3.37; \ d = 1.84 \)

IMPRINT task taxons affected: information processing, numerical analysis
Quantitative framework

- **Proposal:** develop a quantitative framework
  
  Extend current research (with Matt Jones)
  
  Predict training effects in complex military tasks

- **Approach**
  
  Develop theoretical framework based on laws of learning and memory
  
  Encompass multiple training principles through functional parameters

- **Parallel Work**
  
  Purdue Group is testing and extending the framework in parallel with the CU Group.
Core framework

- **Strength of knowledge from past experience**
  
  Activation $a$ is the sum of all prior similar instances $i$ of a function of time $t$

  $s = $ similarity of current task to past training experience

  $d = $ memory decay rate

  $\beta = $ overall learning rate

  $\alpha = $ ease of generalization across knowledge domains

  $$ a = \sum \beta t_i^{-d} e^{-\alpha s_i} $$

- **Behavioral predictions can be derived for…**

  Accuracy

  Response time
Training principles can be incorporated into the framework

- Power law of practice
- Deliberate practice
- Depth of processing
- Contextual reinstatement
- Procedural vs. declarative training
- Instance- vs. rule-based training
- Knowledge seeding
- Spacing of practice
- Power law of forgetting
- Testing retards forgetting
- Generalization depends on similarity
Future work: A plan for framework development

- **Parameter determination**
  
  determine dependencies of parameters on task variables
  
  parameters may be a function of individual differences

- **Modeling**
  
  use simulations and math analysis to evaluate candidate formulations
  
  analyses will feed back to experimental design for additional empirical evidence
  
  simulations of complex, realistic tasks will evaluate model predictions if empirical evidence is not determinative

- **Distinguish…**
  
  empirically grounded variables from ad hoc assumptions
  
  parameter values that are universal, vary by taxon, or vary by task
IMPRINT modeling of RADAR
(Carolyn Buck-Gengler)

- Modeling a militarily-relevant, complex cognitive task in IMPRINT

- Last year – Initial IMPRINT model of an experiment using RADAR task
  (Young, Healy, Gonzalez, & Bourne, 2007)
  - Frame-level analysis of experimental data
  - Modeling response times for visual target detection
  - Modeling detection hit rate
  - Modeling secondary parallel task of tone counting
  - Modeling two sessions - training and (delayed) test

- This year
  - Modeling false alarm rate effects
  - Final adjustment of all parameters
  - Goodness-of-fit analysis
RADAR experiment to model

- **RADAR task** – *scanning visual stimuli for targets*
  - Repeated visual scan of moving stimuli (targets and distractors)
  - Search for pre-assigned target(s) among distractors
  - Manual response on detection

- **Manipulations in this experiment**
  - Variability of task difficulty
    - Processing load – 1 or 4 targets & search items
    - Mapping type – consistent or variable mapping
  - Manipulation of concurrent secondary task of tone counting
  - Training and testing sessions, crossed with presence of 2ndary task

- **Performance assessments** - speed and accuracy
  - Speed of response to targets
  - Accuracy of response (hits and false alarms)
IMPRINT model of RADAR review

- **RT effects – mapping type and processing load**
  - CM responses faster than VM responses
  - Lower memory load (1 target) faster than higher memory load (4 targets)
  - Interaction: VM 4 >> VM 1, CM 4 > CM 1

- **RT effects – tone counting**
  - Tone counting cost (tone presentation & deviant tone incrementing)
  - Training with tone counting incurs cost at delayed test

- **Accuracy effects – Hit Rate**
  - HR lower on VM trials than on CM trials
  - Interaction: VM 4 < VM 1 = CM 4 = CM 1
  - HR lower with tone counting and after training with tone counting

- **Model architecture (stochastic)**
  - Main network: computer presentation of visual stimuli
effects of tone counting on RT
effects of training with tone counting on retention
  - Goal network: subject’s visual search RTs
  - subject’s accuracy (HR, FA rate, FA avoidance learning)
Modeling false alarms

- **False alarm effects** – complex function of mapping type, processing load, session, & tone counting condition

- **General effects**
  - More FAs in VM trials than in CM trials
  - In training, more FAs in VM 4 trials than VM 1 trials
  - More FAs at the beginning of each session
  - More FAs with tone counting
  - Training with tone counting incurs FA cost at delayed test

- **Learning to avoid FAs**
  - Improvement in FA rate seen in all blocks (i.e., all trial types)
  - Large improvement in FA rate when block-initial rate was high:
    (1st block both sessions, VM 4 blocks in training, all VM test blks)

- **Implementation**
  - 1st blocks of sessions numerically specified
  - Non-initial blocks relative to prior block
  - block-initial rates adjusted downward on each trial
Model goodness-of-fit

- Estimate parameter values from data (using block means).
- Limited search of parameter space using model runs.
- Run model twice with final (best) values
- Measure model goodness-of-fit (RMSE and $r^2$)
Model evaluation

- Comparison of model (2 runs) and experimental data
  - RMSE and $r^2$ calculated using (16) block means

<table>
<thead>
<tr>
<th></th>
<th>$r^2$</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Response Time</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model Run 1 vs. Data</td>
<td>0.975</td>
<td>54.2</td>
</tr>
<tr>
<td>Model Run 2 vs. Data</td>
<td>0.984</td>
<td>49.1</td>
</tr>
<tr>
<td>Model Run 1 vs. Model Run 2</td>
<td>0.980</td>
<td>34.3</td>
</tr>
<tr>
<td><strong>Hit Rate</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model Run 1 vs. Data</td>
<td>0.969</td>
<td>0.017</td>
</tr>
<tr>
<td>Model Run 2 vs. Data</td>
<td>0.940</td>
<td>0.025</td>
</tr>
<tr>
<td>Model Run 1 vs. Model Run 2</td>
<td>0.974</td>
<td>0.017</td>
</tr>
<tr>
<td><strong>False Alarm Rate</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model Run 1 vs. Data</td>
<td>0.461</td>
<td>0.062</td>
</tr>
<tr>
<td>Model Run 2 vs. Data</td>
<td>0.285</td>
<td>0.067</td>
</tr>
<tr>
<td>Model Run 1 vs. Model Run 2</td>
<td>0.210</td>
<td>0.081</td>
</tr>
</tbody>
</table>
Summary and future work

- Performance Shaping Functions, Quantitative framework, and IMPRINT Modeling, discussed separately here, are all interrelated, all based on the same test beds and empirical foundation. Progress on any one of them helps to advance or clarify work on the other two.

- Cognitive processes in complex, militarily relevant task (RADAR) were successfully modeled with IMPRINT

- Optimization of full IMPRINT model of RADAR

- Fusion model currently being developed
End