Goals of Project

• Construct a theoretical and empirical framework for training
• Predict the outcomes of different training methods on particular tasks
• Point to ways to optimize training
Three Interrelated Project Components

(1) Experiments and data collection

(2) Taxonomic analysis

(3) Predictive computational models
Organization of Present Meeting

(I) Introduction

(II) Progress in Second Year and Future Plans For Project
   (A) Experiments
   (B) Taxonomy
   (C) Models

(III) Caucuses and Feedback
MURI Personnel

(1) University of Colorado (CU)
   Alice Healy, Principal Investigator
   Lyle Bourne, Co-Principal Investigator
   Bengt Fornberg, Co-Investigator
   Ron Laughery, Consultant
   Bill Raymond, Research Associate
   Carolyn Buck-Gengler, Research Associate

(2) Carnegie Mellon University (CMU)
   Cleotilde Gonzalez, Co-Investigator

(3) Colorado State University (CSU)
   Ben Clegg, Co-Investigator
   Eric Heggestad, Co-Investigator

(4) Purdue University (Purdue)
   Robert Proctor, Co-Investigator
Meeting Presenters

(1) Overview and Coordinate
   Healy & Bourne

(2) Experiments
   (a) Development & Testing of Training Principles
       Healy & Bourne
   (b) Acquisition & Retention of Basic Components of Skill
       Proctor
   (c) Levels of Automation, Individual Differences, & Team Performance
       Clegg & Heggestad

(3) Taxonomy
   Raymond

(4) Models
   (a) ACT-R
       Gonzalez
   (b) IMPRINT
       Buck-Gengler
   (c) Model Assessment
       Fornberg
Positive Committee Comments

(1) The government committee was pleased with the scientific content of most of the presentations at the annual review. Cooperation among the MURI team members has been good.

(2) The committee was glad to see the excellent scientific progress made by the MURI team during the year.

(3) The planning and coordination of the projects is at a high level. The projects mutually reinforce each other and create a combined effort greater than the sum of the individual projects.

(4) The work on low-level tasks has produced excellent results, including identification of training principles not previously formulated.

(5) The ramp-up in the effort on IMPRINT recommended in the report on the kickoff meeting has been carried out and results indicating excellent progress were presented.

(6) The work on taxonomy previously requested by the committee, has proceeded well.

(7) It is a positive factor that individual differences are being taken into account in part of the research.
Key Critical Committee Comments

(1) The work needs to ramp up from simple tasks to more complicated tasks close to those for which the Army trains.
(2) The medium of delivery in tasks and in training is important. A number of major new options that are currently being proposed to DoD need to be considered.
(3) The detailed investigations that the project “Purdue Experiments” is carrying out have potentially large impact on the other projects of the MURI. At present, however, the potential is not being realized. This project should be integrated with the other projects.
(4) The project “Model Evaluation” is far behind a reasonable schedule and has no direct connection with any of the other projects of the MURI. All projects that have modeling as a component need to be connected and integrated.
(5) More information on the team’s goal at the end of five years and how the team is going to reach that goal is needed.
Key Comments from Government Committee

(1) The work needs to ramp up from simple tasks to more complicated tasks close to those for which the Army trains. Scientific feasibility needs to be kept in mind to guide this ramp-up process--reliable information on tasks of medium complexity may be more important than less reliable information on tasks of large complexity.

(1) At the 2006 Annual Meeting, we emphasized a single task (data entry) across experiments, taxonomy, and models, so that we could show the coherence of our work. However, we study, in all phases of the project, tasks considerably more complex than data entry, including, for example, target detection and decision making, message comprehension in a navigation task, and information lookup in a data base.
Key Comments from Government Committee

(2) The results presented on automation (CSU) with awareness of individual differences were good and can lead to increased understanding of the role of the medium of delivery. The medium of delivery in tasks and in training is important. A number of major new options that are currently being proposed to DoD need to be considered.

(2) We explored some of these major new options during a site visit to the SSRU. These options will be included in the taxonomy and will influence some of our empirical work. Particularly important are training simulations of battlefield deployment and control of unmanned aerial vehicles. In addition, our recent and continuing research has explored differences in learning as a function of modality of presentation.
Key Comments from Government Committee

(3) The project “Purdue Experiments” has potentially large impact on the other projects of the MURI and those other projects have the potential for impacting this project by providing guidance for what questions need to be answered. At present, however, the potential is not being realized.

(3) To integrate the Purdue research with the other projects at CU, monthly telephone meetings were held at which our research was coordinated, and the Purdue team conducted an experiment that incorporated the main focus of the Purdue research, stimulus-response compatibility, into one of the tasks, data entry, that has been studied extensively by the CU team and modeled using ACT-R and IMPRINT.
Key Comments from Government Committee

(4) The project “Model Evaluation” is far behind a reasonable schedule and has no direct connection with any of the other projects of the MURI. All projects that have modeling as a component need to be connected and integrated.

(4) Fornberg has re-focused the modeling evaluation effort from exploring general opportunities to actual measurements. Evaluation of the ACT-R and IMPRINT models is based on (a) goodness of the fit between experimental observations and output from the simulations, (b) mathematical tractability, (c) computational efficiency, and (d) scalability. Fornberg has written a Matlab implementation of the algorithms used in the IMPRINT model. This Matlab model yields a benchmark for computing time, an index of computational efficiency.
Key Comments from Government Committee

(5) More information on the team’s goal at the end of five years and how the team is going to reach that goal is needed. How do intermediate goals planned for years 2 and 3 provide a path to the goal at the end of 5 years. At the 2-year annual meeting, an estimate of the final goal product of this MURI should be stated and the path from here to there should be outlined.

(5) We provide in the following: (a) a specification of goals for the end of the MURI grant and (b) a meta plan of how to get from where we are now to those goals.
Goals for the End of the MURI Grant

1. We will produce a “manual” of empirically-based training principles.
   a. The manual will contain instructions for the use of principles by military trainers.
   b. The principles will have an empirical basis deriving from a large set of laboratory experiments including those involving tasks of sufficient complexity to ensure their applicability to military training.
Goals for the End of the MURI Grant

2. We will develop a taxonomy of task type, training method, and performance context that will combine with performance measures and empirically-based training principles to form a space of training effects.

a. The training effects space will include quantitative functions (i.e., performance shaping functions) for critical principles.

b. The task taxonomy will be capable of being mapped to IMPRINT task taxons and workload values.

c. The taxonomy should be useful as a research planning tool as well as a guide for predictive modeling using either ACT-R or IMPRINT.
Goals for the End of the MURI Grant

3. We will produce an integrated set of ACT-R cognitive models.
   a. The models will represent the implementation of multiple tasks of different degrees of complexity.
   b. The models will involve multiple training principles and training effects.
   c. We will provide an easy way to manipulate some training and ACT-R parameters in the models to produce predictions of the effects of those parameters on various performance measures.
Goals for the End of the MURI Grant

4. We will produce specific models in IMPRINT for data entry, RADAR, and information integration, demonstrating that IMPRINT can be used as a modeling tool for laboratory tasks.

5. We will provide a comparative evaluation of ACT-R and IMPRINT models for a small number of specific experimental tasks under defined conditions (i.e., data entry, RADAR, and information integration). This comparison will address issues of mathematical and computing tractability and scalability within given task demands and contexts.
Meta Plan

We get from where we are now to the deliverables just outlined by following the plan outlined in our grant proposal. We assume that it is sufficient for us to stay within the framework outlined in our approved and funded grant proposal. However, it should be noted that in addition to what we promised to accomplish in the proposal, we will be undertaking some other work suggested in the course of the review of our project (e.g., the planning matrix, the development of new, more complex experimental tasks of military relevance for study in the laboratory, and intensive collaborative research among MURI team members).
Significant Meetings Related to Army Training over Last Year

(1) Healy, Army Science of Learning Workshop
    August 2006, Hampton

(2) Bourne & Raymond, SSRU
    February 2007, Orlando

(3) Goldberg, site visit
    February 2007, Boulder

(4) Raymond, IMPRINT Technical Interchange
    April 2007, Boulder
Development and Testing of Training Principles: Completed Experiments

(1) Tests of the generality across tasks of individual principles

(2) Tests of multiple principles in a single task

(3) Tests of principles in complex, dynamic environments

(4) Developing and testing new principles
Cognitive Antidote Principle

Adding cognitive complications to a routine task overcomes decline in accuracy due to fatigue.
Design

- Subjects typed 960 unique 4-digit numbers divided into 15 blocks of 64 numbers, with no rest between blocks, using their preferred (right) hand.
- Four groups of subjects (8 per group):
  - ALTERNATING (concluding keystroke alternated between + and -)
  - RELATIVE MAGNITUDE (concluding keystroke + or - depending on the relative magnitude of the first and second pairs of digits)
  - + CONTROL (concluding keystroke always +)
  - - CONTROL (concluding keystroke always -)
Control (+)
Control (-)
Relative Magnitude
Alternating
Total Response Time (in log ms)

First  Second  Third
1  2  3  4  5  1  2  3  4  5  1  2  3  4  5
3.30
3.35
3.40
3.45
3.50
3.55
3.60

- Control (+)
- Control (-)
- Relative Magnitude
- Alternating
Training Difficulty Principle

Any condition that causes difficulty during learning may facilitate later retention and transfer.
Design

Between-Subjects Variables

Tone Condition Week 1 (Tone, Silent)
Firing Response Condition Week 1 (Yes, No)
Tone Condition Week 2 (Tone, Silent)
Firing Response Condition Week 2 (Yes, No)

Half of the subjects same tone condition each week; half different conditions

Half of the subjects same firing response condition each week; half different conditions

Within-Subjects Variables

Block (first, second)
Week (first, second)
Proportion of False Alarms

<table>
<thead>
<tr>
<th>Block</th>
<th>Training 1</th>
<th>Training 2</th>
<th>Testing 1</th>
<th>Testing 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train Fire</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Train No Fire</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The graphs illustrate the proportion of hits for different sessions and conditions. The x-axis represents the session (Training or Testing), and the y-axis represents the proportion of hits.

In the top graph, the bars show the proportion of hits for Train Fire and Train No Fire conditions, with error bars indicating variability. The bottom graph shows the d' values for the same conditions, with similar error bars.

The legend indicates the different conditions:
- Red: Train Fire
- Blue: Train No Fire
Training Difficulty
Any condition that causes difficulty during learning may facilitate later retention and transfer.

Specificity of Training
Retention and transfer are depressed when conditions of learning differ from those during subsequent testing.

Variability of Practice
Variable practice conditions typically yield larger transfer effects compared with constant practice conditions.
Left 2 squares
Down 2 levels
Forward 1 step
Design

Between-Subjects Variable
  Training Condition (Easy, Hard, Mixed)

Within-Subjects Variables (at Test)
  Block (1-6)
  Message Length (1-6)

Dependent Measure
  Manual Movement Accuracy
  Oral Readback Accuracy
Strategic-Use-of-Knowledge Principle

Learning and memory are facilitated whenever pre-existing knowledge can be employed, possibly as a mediator, in the process of acquisition.
baby care
bakery
baking supplies
beverages
breakfast goods
canned goods
cereal
cleaning
condiments
dairy
deli
desserts
dry goods
frozen food
fruits
health foods
meats
paper products
personal hygiene
picnic
snacks
soft drinks
spices
vegetables
almonds
Design

Within-Subjects
- Phase (training, testing)
- Block within phase (1-4)
- Item difficulty (easy, difficult)

Between-Subjects
- Content type of first list
  (grocery, department)
- Category organization of first list
  (alphabetical or conceptual)
- Same or different content between phases
- Same or different organization between phases
- List order (1, 2 or 2, 1)
Testing

RT per click (in s)

<table>
<thead>
<tr>
<th>Block</th>
<th>Same Content</th>
<th>Different Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.6</td>
<td>1.4</td>
</tr>
<tr>
<td>2</td>
<td>1.2</td>
<td>1.0</td>
</tr>
<tr>
<td>3</td>
<td>1.0</td>
<td>0.8</td>
</tr>
<tr>
<td>4</td>
<td>0.8</td>
<td>0.6</td>
</tr>
</tbody>
</table>
Memory Constriction Principle

The time span from which memories can be retrieved shrinks as stress increases
Design

24 CU undergraduates (12 each group)
2 x 2 x 2 mixed factorial ANOVA for each measure (test trials only)
– Between: Time pressure (no, yes)
– Within:
  Memory type (retrospective, prospective)
  Retention Interval (short term, long term)

Dependent measure: Accuracy
– Association memory
  • Correct color regardless of location
– Memory of intention to do something
  • Correct location regardless of color
Serial Position Principle

Retention is best for items at the start of a list (primacy advantage) and at the end of a list (recency advantage).
Design

- 48 trials
- 7-item sequences
- 10 x 10 grid
- 3 x 3 damage radius
- One optimal firing location
- Sequential and simultaneous presentation
- Free, serial, and no recall
- Scoring:
  - 100% for a direct hit
  - 55% for adjacent square, sharing a side
  - 20% for adjacent square, sharing a corner
Imagine you are a remote gunner in the army. This means your job is to monitor the battlefield from miles away, controlling anti-infantry missiles.

You will see a green “x” in the center of a grid representing a battlefield. The red blips you will see represent enemy soldiers. These blips will appear one at a time on the screen.

Your job is to pay close attention to the exact locations of enemies as they are presented on the screen. You must immediately make a decision of where to fire in order to damage as many enemies as you can, to the greatest possible extent. Using a mouse, you will choose a firing location and click when you wish to fire. Each battlefield situation will have a single ideal firing location. Maximum damage is inflicted when the ideal square is fired upon.

Some damage is inflicted whenever a missile is fired at a square adjacent to an enemy, but direct hits inflict maximum damage to a given enemy. Hitting adjacent squares that share entire sides with an enemy inflicts more damage than does hitting squares that share just a corner with an enemy. The ideal firing location may sometimes be an empty square, if it is adjacent to a number of enemies.

These are state-of-the-art weapons, so they always hit where you aim. You can only fire one missile per set of enemy locations.

Immediately after firing, you will see a screen with a bull’s-eye showing the ideal firing location for the previous set of enemy locations, your score on that shot, and where you fired.

You will then be asked to click on the location of each of the enemies, in the order that they appeared.

Click Start when ready to start experiment. It will begin immediately.
Key Critical Committee Comments

(1) The work needs to ramp up from simple tasks to more complicated tasks close to those for which the Army trains.

(2) The medium of delivery in tasks and in training is important. A number of major new options that are currently being proposed to DoD need to be considered.

(3) The detailed investigations that the project “Purdue Experiments” is carrying out have potentially large impact on the other projects of the MURI. At present, however, the potential is not being realized. This project should be integrated with the other projects.

(4) The project “Model Evaluation” is far behind a reasonable schedule and has no direct connection with any of the other projects of the MURI. All projects that have modeling as a component need to be connected and integrated.

(5) More information on the team’s goal at the end of five years and how the team is going to reach that goal is needed.
What Can We Promise at the End?

(1) We will produce a “manual” of empirically-based training principles.

(2) We will develop a taxonomy of task type, training method, and performance context that will combine with performance measures and empirically-based training principles as additional (orthogonal) categorical dimensions to form a space of training effects.

(3) We will produce an integrated set of ACT-R cognitive models.

(4) We will produce specific models in IMPRINT for data entry, and RADAR, demonstrating that IMPRINT can be used as a modeling tool for laboratory tasks.

(5) We will provide a comparative evaluation of ACT-R and IMPRINT models for a small number of specific experimental tasks under defined conditions (i.e., data entry and RADAR).