U.S. Army Research, Development and Engineering Command

Providing Agility in C2 Environments Through Networked Information Processing: A Model of Expertise

TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

Kevin Chan, Jin-Hee Cho, US Army Research Laboratory
Sibel Adali, Rensselaer Polytechnic Institute
Jennifer Mangels, Olta Hoxha, Ksenia Lila, CUNY Baruch College
Damon Abraham, CUNY Graduate Center

ICCRTS 2014
Track # 4
Paper # 33
June 17, 2014

Problem

- Decision making in command and control (C2) environments requires gathering, processing and sharing of all available information to establish intelligence and to increase situation awareness.
- Complex criteria play a role in such a decision:
  - trustworthiness (reliability, benevolence) of the sources
  - ability (expertise) of analysts to analyze and filter the information
  - underlying difficulty/complexity of the decision space
  - other network level factors like social influence

US ARMY RDECOM

Contribution

ARL

- Summarized experiments by CUNY that study the expertise of individuals who answer general knowledge domain questions in an experimental setting
- Studied the expertise and difficulty component of this process
- Incorporated an initial mathematical model of expertise and difficulty into an agent model for information processing in networks

Connections to C2 and ELICIT

Command and Control (C2)

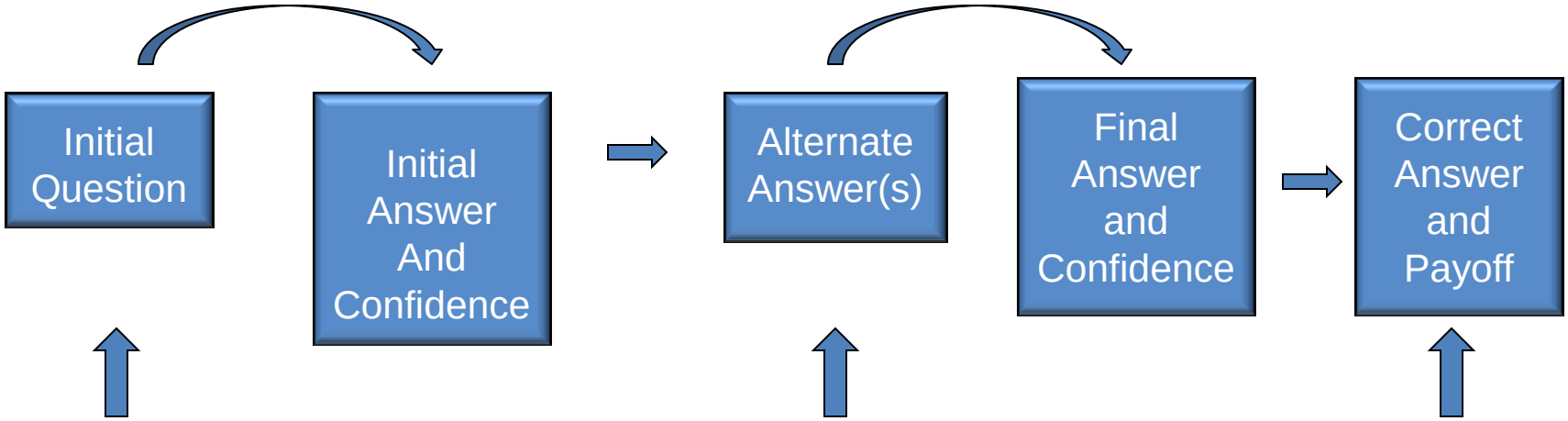
- Developed distributed information sharing framework that enhances situation awareness
- Predicted correctness of information sharing, given specific level of problem difficulty

Agent Models

- Developed ELICIT- based and other agent based models for information sharing scenario
- Predicted correctness of information processing, given various levels of problem difficulty

Experiment Flow

- Set of 120 questions, titrated to achieve initial accuracy of .33



Test	Treatment	Abbreviation	# of subjects
(1-a)	One alternate answer with no time pressure	*alt1off*	33
(1-b)	One alternate answer with time pressure	*alt1on*	39
(3-a)	Three alternate answers with no time pressure	*alt3off*	39
(3-b)	Three alternate answers with time pressure	*alt3on*	37

Table 1: Description of experiment treatments and number of subjects tested.

- Cleaning of data, titration breakers, confused participants

Experiment Setup

Demographics

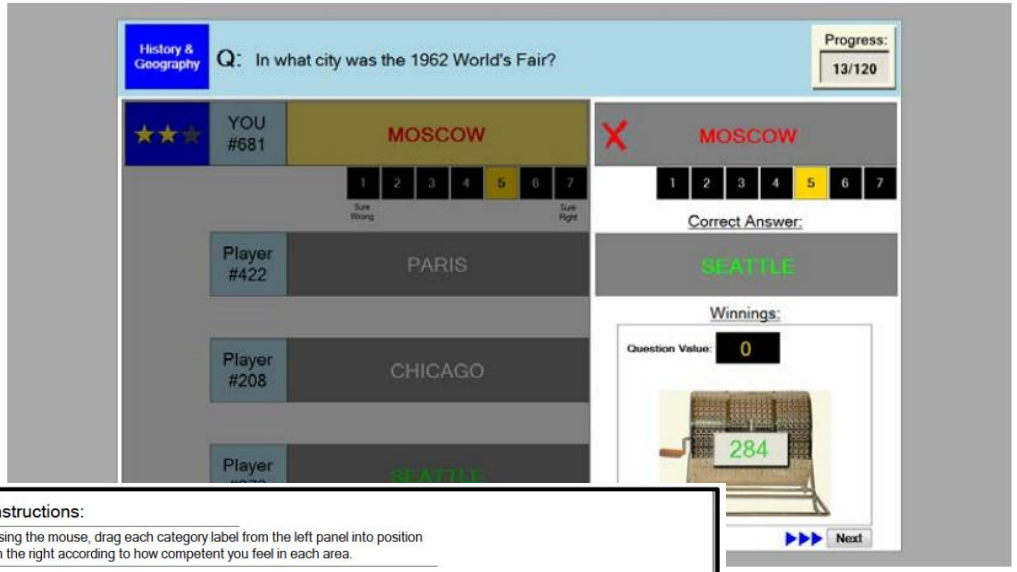
- College students
- Expertise of domains

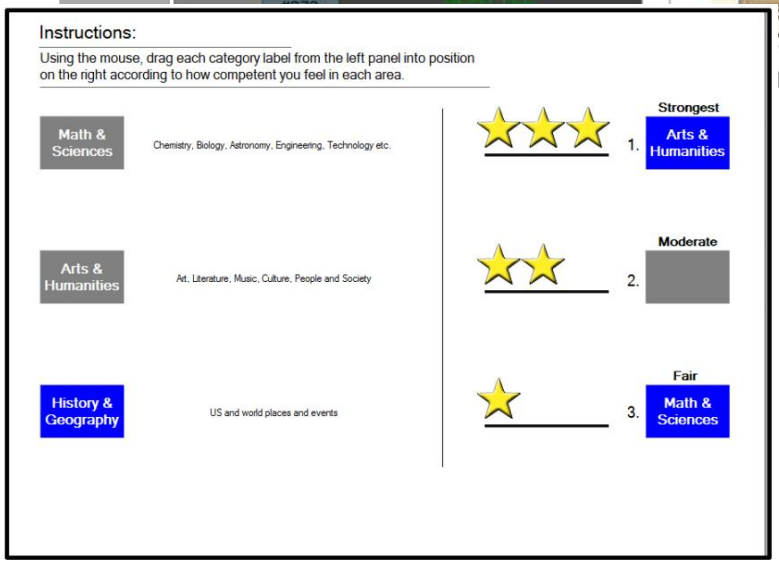
Metrics

- Initial and Final Accuracy
- Initial and Final Confidence

Treatments

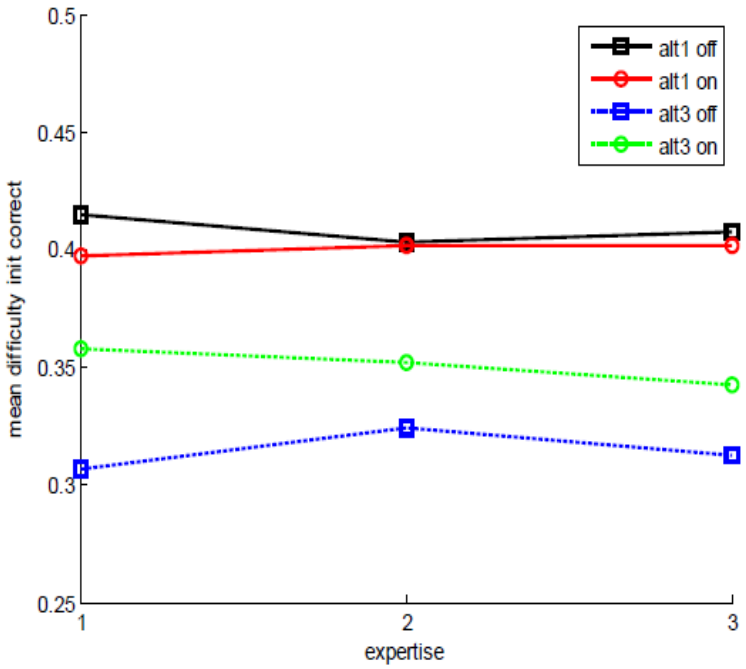
- Alternate answers: {1, 3}
- Time pressure: {on, off}

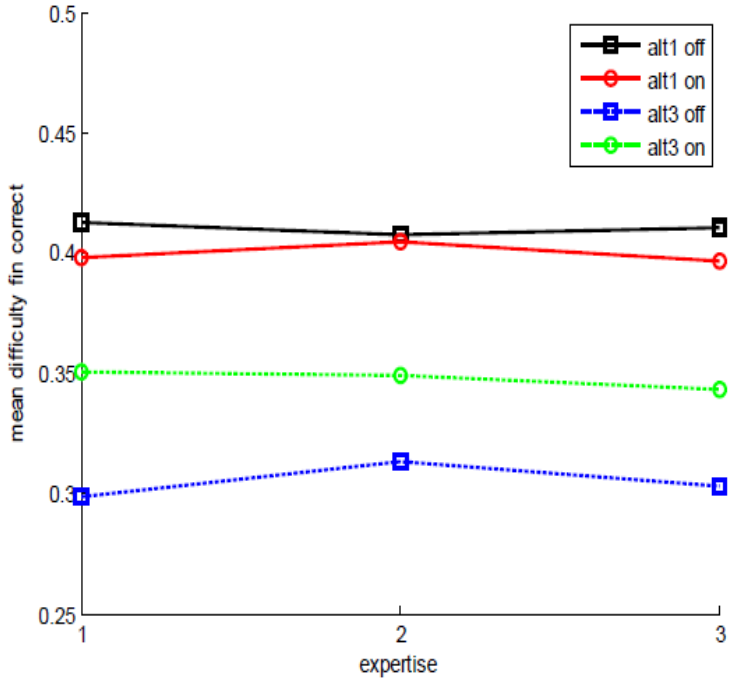




Correctness vs. Expertise

- Mean difficulty of correct accuracy vs. expertise for each treatment





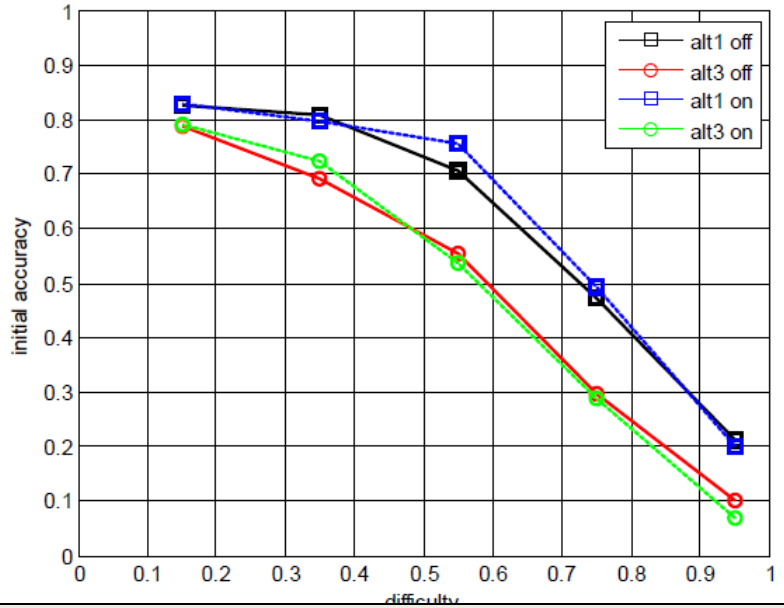
Correct answers don't exhibit significant differences when it comes to self-reported expertise

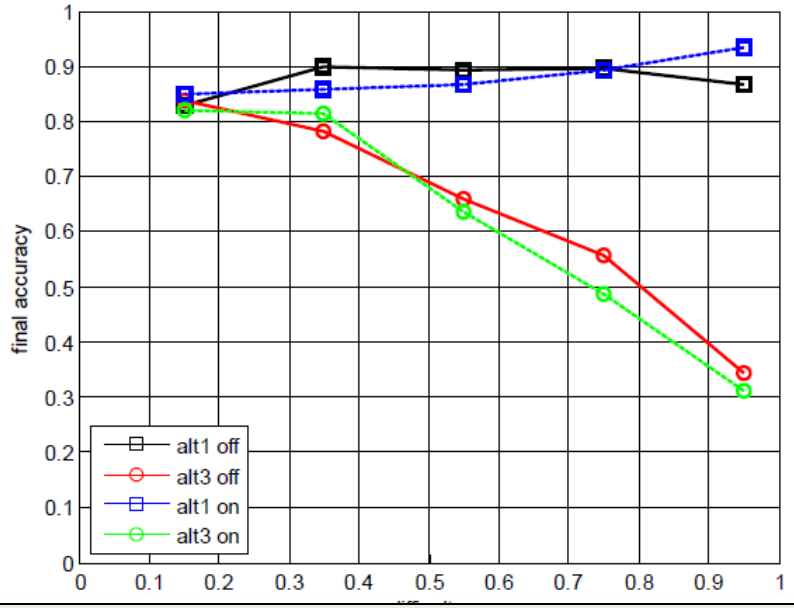
US ARMY RDECOM

Accuracy vs. Problem Difficulty

ARL

- Initial and Final accuracy versus problem difficulty
- Data Cleaning: Initial accuracy (removed wrong|wrong for alt1)





Based on the number of alternate choices vs. problem difficulty, we see an increase in accuracy

Statistical Analysis

- Kolmogorov-Smirnov test (K-S test) determines if two empirical distributions are from the same distribution
- CDFs, $F_1(x)$ and $F_2(x)$

$$D = sup_x |F_1(x) - F_2(x)|$$

- Q_{KS} is a measure of the probability that the distributions are the same

$$\Pr(d > \text{observed}) = Q_{KS}\left(D\left(\sqrt{\frac{n_1 + n_2}{n_1 n_2}} + .12 + .11/\sqrt{\frac{n_1 + n_2}{n_1 n_2}} \right) \right)$$

Statistics between Treatments

Expected (lack of) significance for initial accuracy, difference in final accuracy

- Initial Accuracy (less significant difference in performance)

	alt1off	*alt1on*	*alt3off*	*alt3on*
alt1off		.0535	.2306	.2264
alt1on			.2220	.2212
alt3off				.0309
alt3on				

- Final Accuracy (significant difference in performance)

	alt1off	*alt1on*	*alt3off*	*alt3on*
alt1off		.0999	.4577	.4769
alt1on			.4865	.5054
alt3off				.0546
alt3on				

Statistics between Expertise

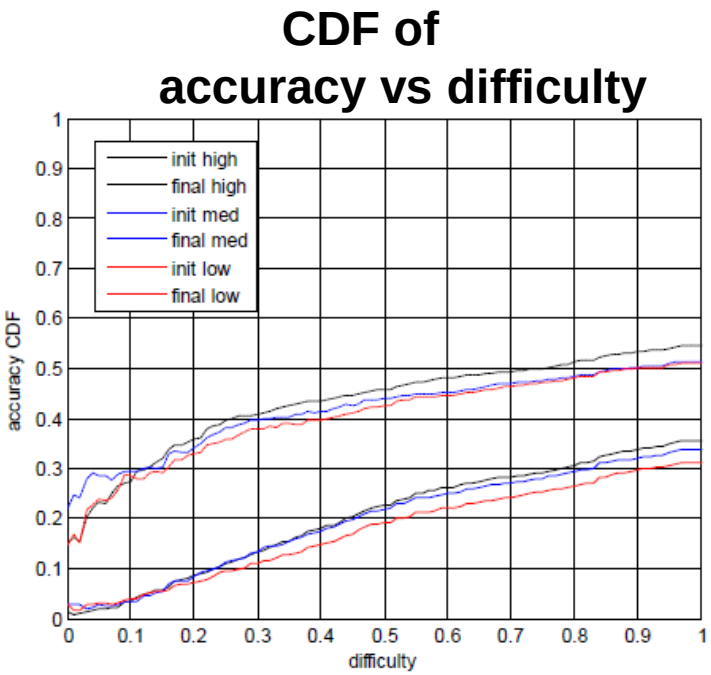
Minor statistical significance between accuracy for different expertise

- Initial Accuracy (little statistical significance)

	High exp	Med exp	Low exp
High exp		.0167	.0359
Med exp			.0266

- Final Accuracy (little statistical significance)

	High exp	Med exp	Low exp
High exp		.0715	.0354
Med exp			.0751



Experiment Conclusions

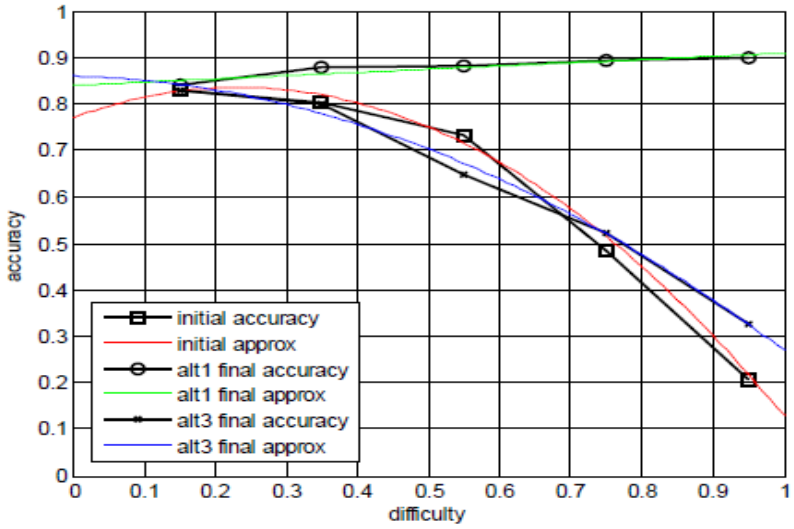
- Significance is observed:
  - when comparing the *alt1* and *alt3* cases
  - between the initial and final accuracy

- No significance is observed with respect to the function of time pressure and expertise

Mathematical Formulation

- Regression of initial and final accuracy can inform agent models of accuracy given problem difficulty

Data	Regression	R
Initial accuracy	$-1.2d^2 + .56d + .77$	.034
alt1 final accuracy	$.069d + .84$	.01
alt3 final accuracy	$-.55d^2 - .04d + .86$	



Item Response Theory (IRT)

- Use IRT to understand how problem difficulty can predict accuracy of responses to specific questions

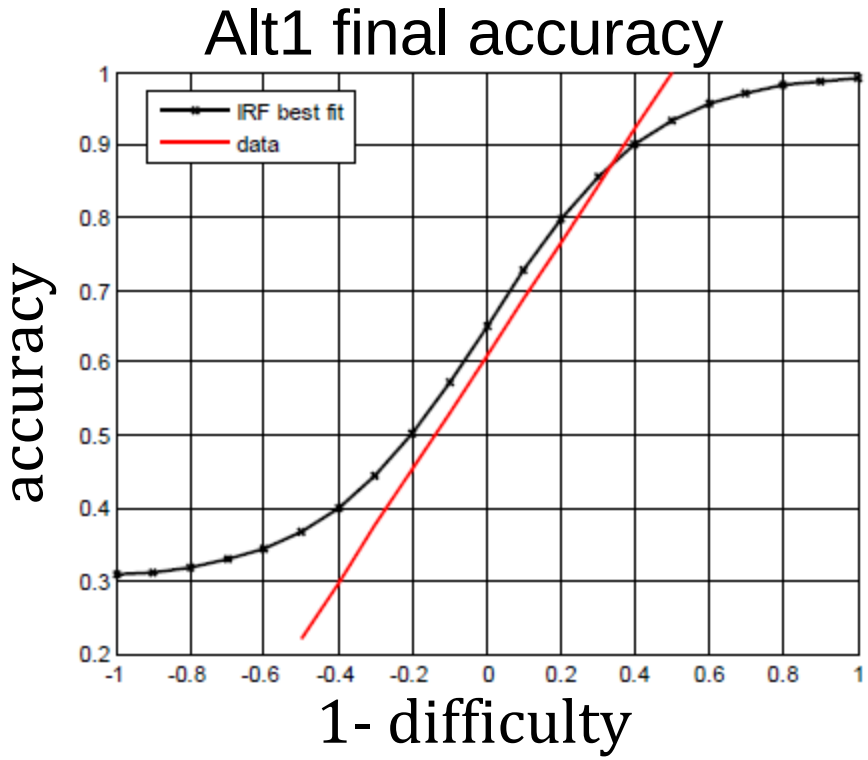
- Three-parameter logistic (*3PL*) model

$$p_i(x) = c_i + \frac{1 - c_i}{1 + e^{-a_i(x - b_i)}}$$

b – difficulty

a – discrimination, slope

c – pseudo-guessing



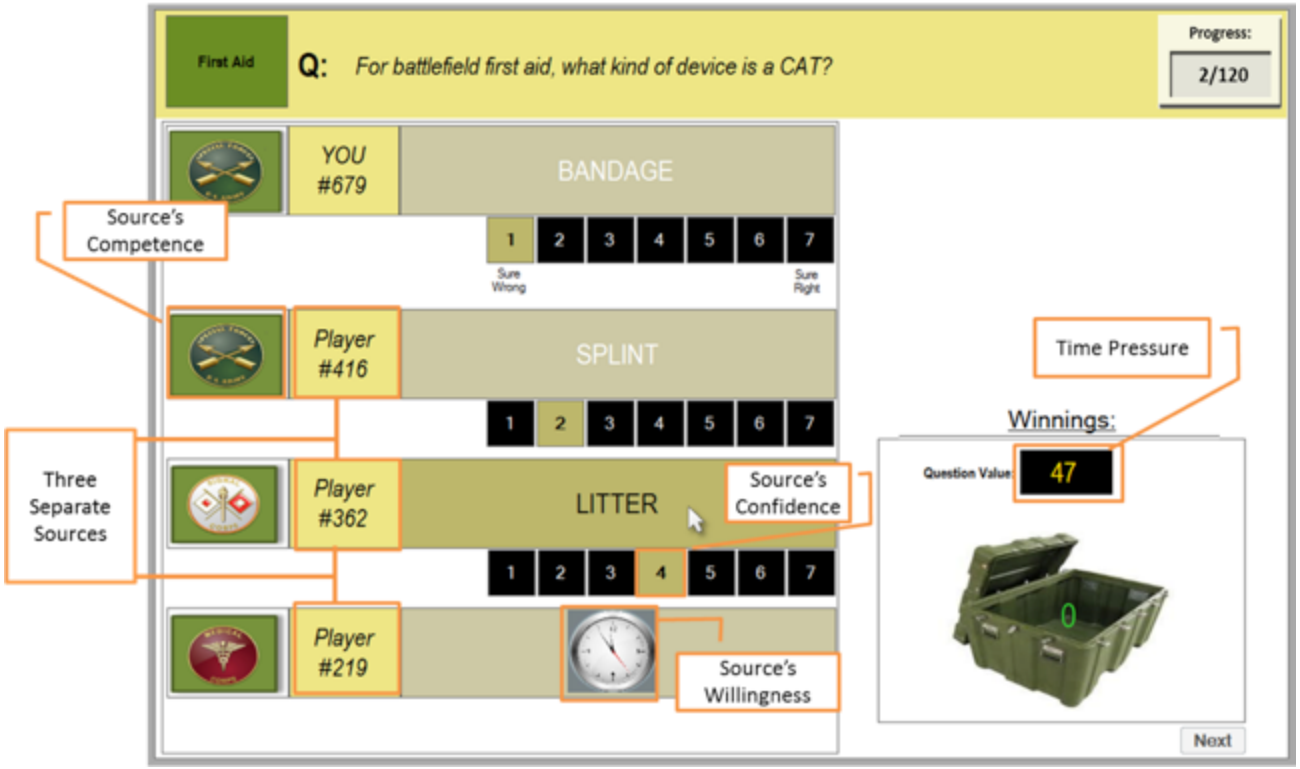
Conclusions

- Expertise / training impacts decision making ability; however, observations do not show any statistical significance in decision making performance
- Problem difficulty vs. decision making performance shows statistical significance
- Mathematical model based on experimental parameters provides capability to configure software agents to match the trends observed in the human experiments

UNCLASSIFIED / APPROVED FOR PUBLIC RELEASE

Military Relevant Experiments

- Create a military relevant study
- Compare against existing normative study



Thank You!

Contact us at:

Kevin Chan

U.S. Army Research Laboratory

kevin.s.chan.civ@mail.mil