



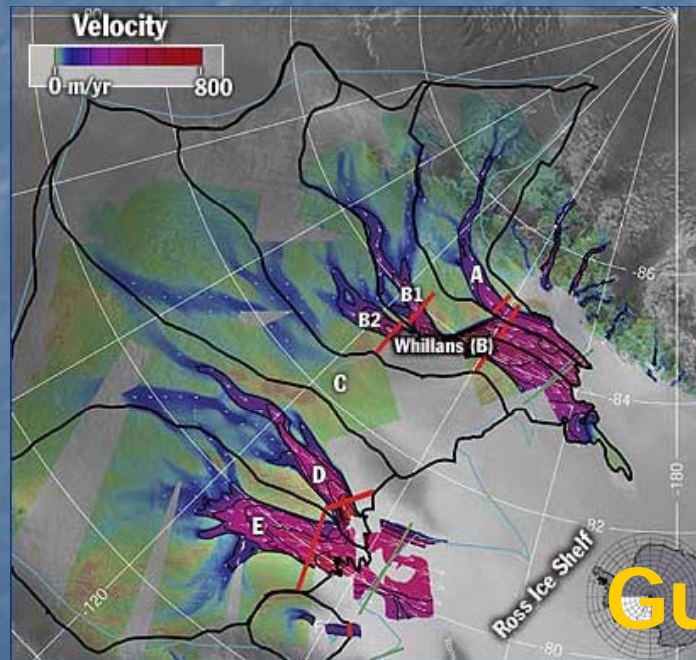
Low-fidelity Simulations For. Assessment

- Joseph Psotka
- Peter Legree, Tiffany Bludau
- Jim Belanich
- ARI
- ICT Oct. 4, 2008

ZPD: Zone of Proximal Development



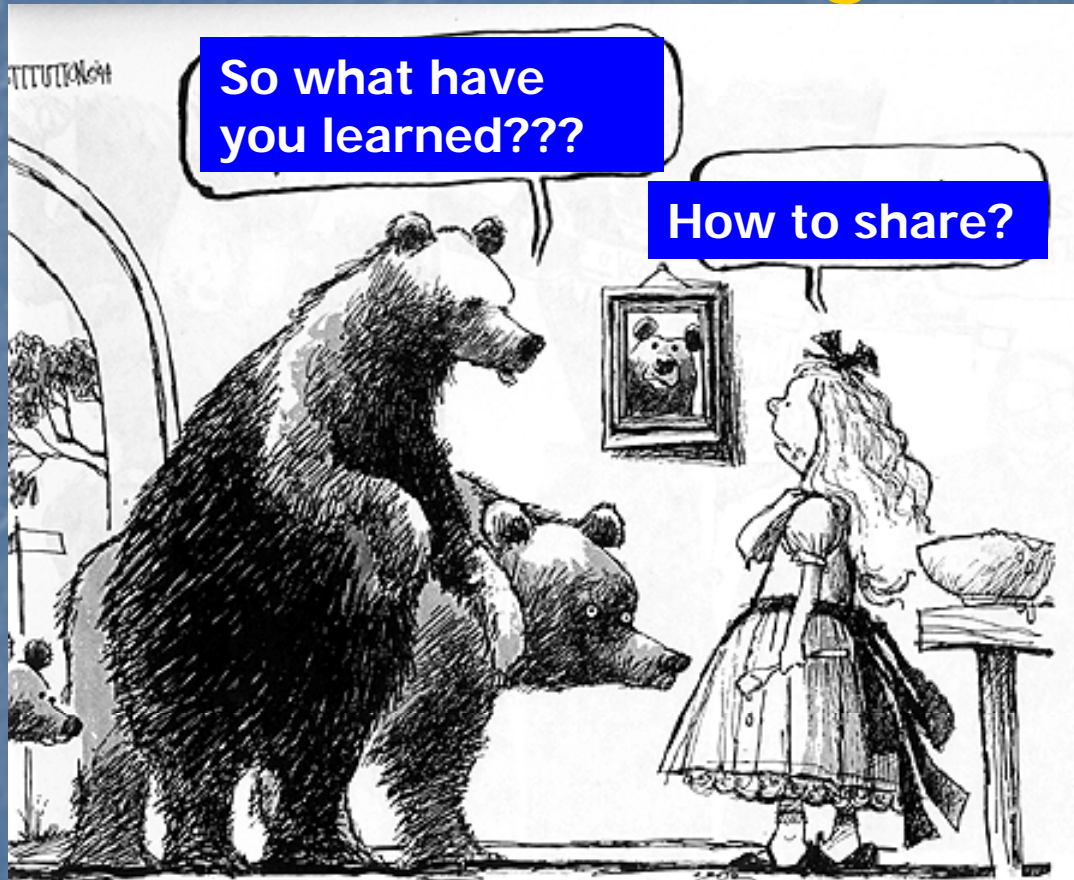
The Growing Edge of Change
is where Development can
be accelerated ...



Guided and
scaffolded ...

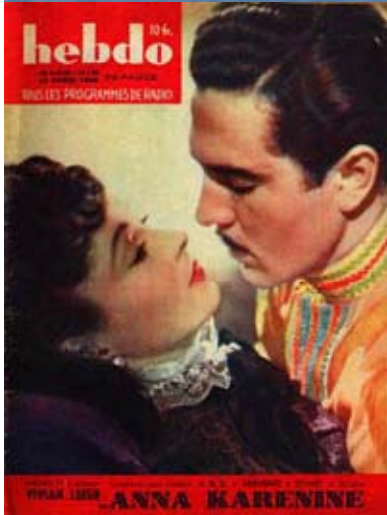
Goldilocks

Neither too advanced,
Nor too elementary;
Just right!



Tolstoy

Happy families are all alike; Every unhappy family is unhappy in its own way



Many diverse opinions can be united into a better consensual whole ..many novices can combine to outperform an expert



It takes a village to grow a child



Oct. 2, 2008

Army Research Institute

4

LTC Dongle is meeting Achmed for the first time, and they are getting to know each other. Dongle knows that Achmed, a Shia, is an important busi

Situation Judgment Tests: Low - Level Simulations



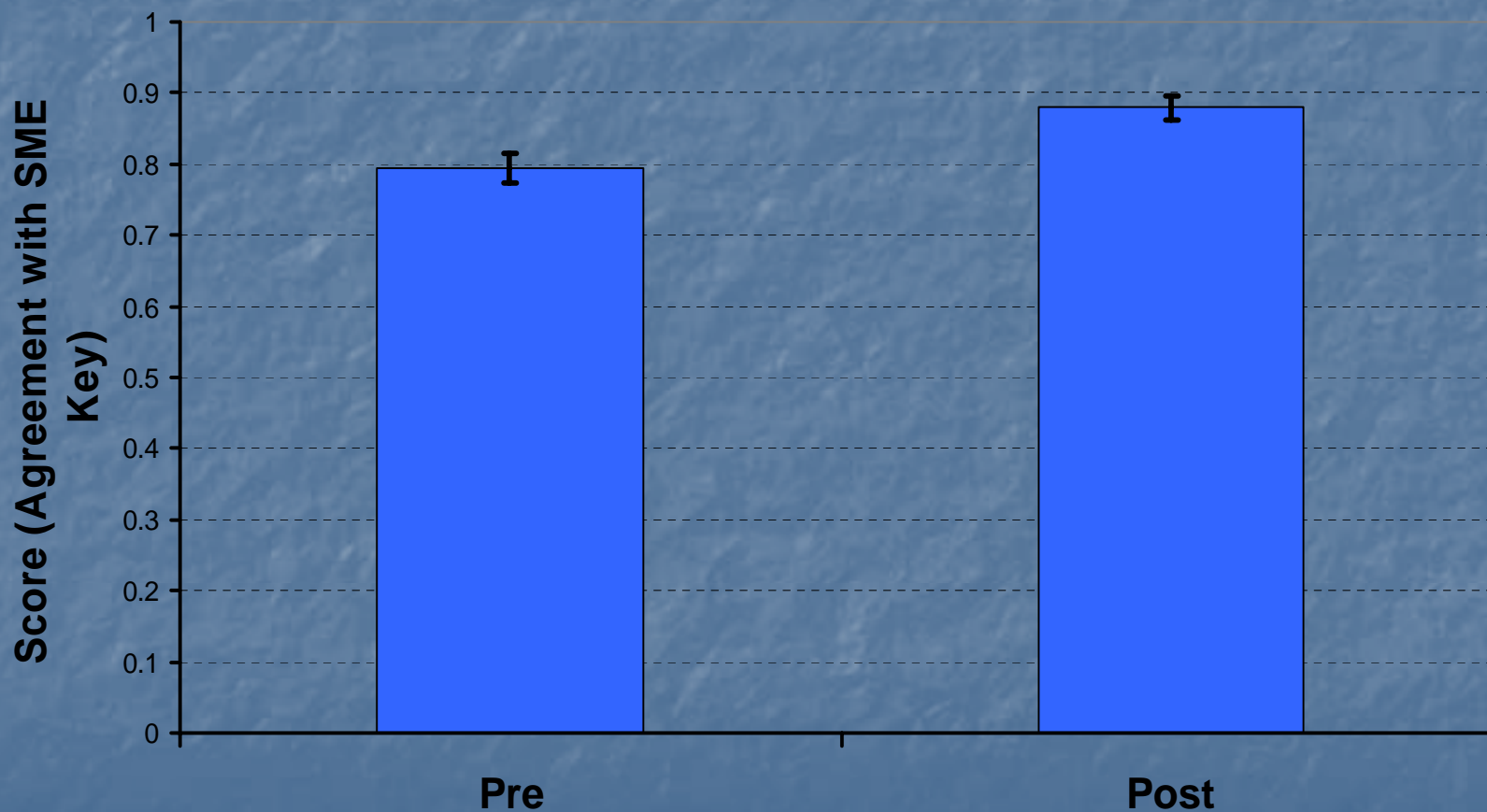
LTC Dongle is meeting Achmed for the first time, and they are getting to know each other. Dongle knows that Achmed, a Shia, is an important businessman with influence in this primarily Shiite area. To develop rapport with Achmed, what sorts of topics of conversation would be good for LTC Dongle to bring up during the meeting?



How much authority he has as a LTC in the US Army.	Poor			Moderate					Good		
	0	1	2	3	4	5	6	7	8	9	10
How is one of the main reasons why there is still unrest in .	Poor			Moderate					Good		
	0	1	2	3	4	5	6	7	8	9	10
How Achmed’s family is doing these days.	Poor			Moderate					Good		
	0	1	2	3	4	5	6	7	8	9	10
How much he enjoys the local food in this area.	Poor			Moderate					Good		
	0	1	2	3	4	5	6	7	8	9	10



Average Scores on Situational Judgement Test (N=11)



Significant difference between pre- and post-test scores (paired samples t-test, $df=10$, $t=-3.52$, $p=.0027$)

Initial Results for ELECT BiLAT Instruction



- Instruction includes:
 - Short lecture/presentation from course instructor (i.e., how to use BiLAT and goals for lesson)
 - About 1 hour to use BiLAT (preparing and conducting simulated meetings with automated tutor/coach feedback)
- Pre/Post-test (N=11)
 - Situational Judgment Test (30 items) used to assess decision-making in bilateral negotiation situations.
 - Agreement scores derived from correlations between student choices and SME scoring key.

Results

- Average pre-test agreement score = 0.79, average post-test agreement score = 0.88
- Significant increase from pre- to post-test scores (*paired samples t-test*, $t(10)=3.52$, $p=.0028$)

TKML Vignettes



- Tacit Knowledge of Military Leadership (TKML)
 - Developed by ARI and Yale University
 - Focus on Interpersonal Leadership skills
 - Validated for Platoon, Company, and Battalion
- A platoon-level scenario:
 - You are a new platoon leader who takes charge of your platoon when they return from a lengthy combat deployment. All members of the platoon are war veterans, but you did not serve in the conflict. In addition, you failed to graduate from Ranger School. You are concerned about building credibility with your soldiers. What should you do?

Knowledge Post: An On - Line Collaborative Learning Environment



Knowledge Post is a standard threaded discussion environment that has been enhanced with LSA.

- Read notes including vignette description
- Write notes in response to those scenarios or prompts.
- Respond to vignette and notes of others
- Search for semantically similar notes
- Receive feedback on your contributions by the *Intelligent Essay Assessor*

Related Notes



KNOWLEDGE POST

[Main Index](#) | [Search](#) | [Messages](#) | [Who's Online](#) | [Profile](#) | [Logout](#) | [Help](#)

[Riley1](#) >> [Trouble in McLouth](#)

Notes related to "Where's PAO?"

Subject	Similarity (0-100)	Find Related		Author	Date
get out of the way	(55)	Notes	References	leaderR8	06/03/02 09:17 AM
FINAL THOUGHTS		Notes	References	leaderR5	06/03/02 09:34 AM
Final METT-T		Notes	References	leaderR7	06/03/02 09:33 AM
Final Thoughts.		Notes	References	leaderR8	06/03/02 09:36 AM
re: First things First	(42)	Notes	References	leaderR8	06/03/02 09:30 AM

Halt all the convoy serials that have not entered the congested areas and try to re-route those to the BSA to support the Brigade.

leaderR2
06/03/02 09:12
AM

Where's PAO?

[Reply](#)

Find Related

[Notes](#)

Related Library

[References](#)

First thing we need here at this scene is the PAO office with MP assistance standing by. Inform the commander of the following serials of the situation and get the company commander working on an alternate route. This will not be over quickly and we don't want the whole support unit stuck here.

KP vs. Paper & Pencil



- Collected responses from over 200 officers at different posts
- Officers' responses graded by four military experts
- Higher quality responses using KP
- Demonstrable learning using KP

TLAC On-line:

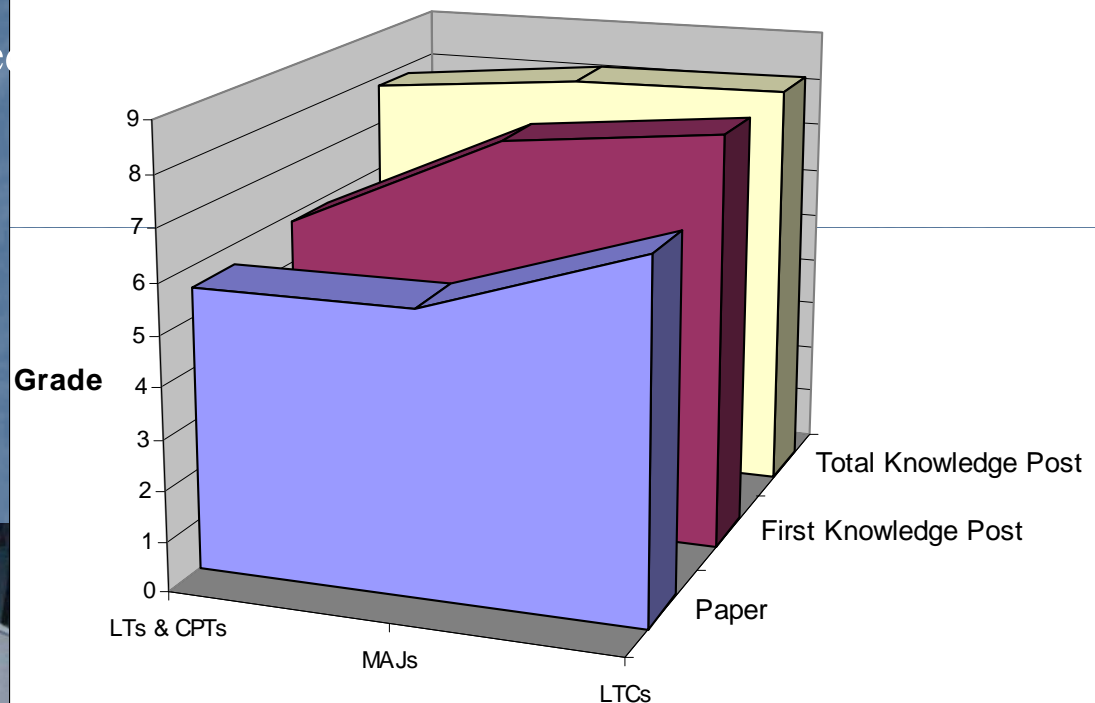
Paper vs Knowledge Post for TLAC Across Ranks



Two forces appear to drive the superior performance and greater learning evidence in the online environment:

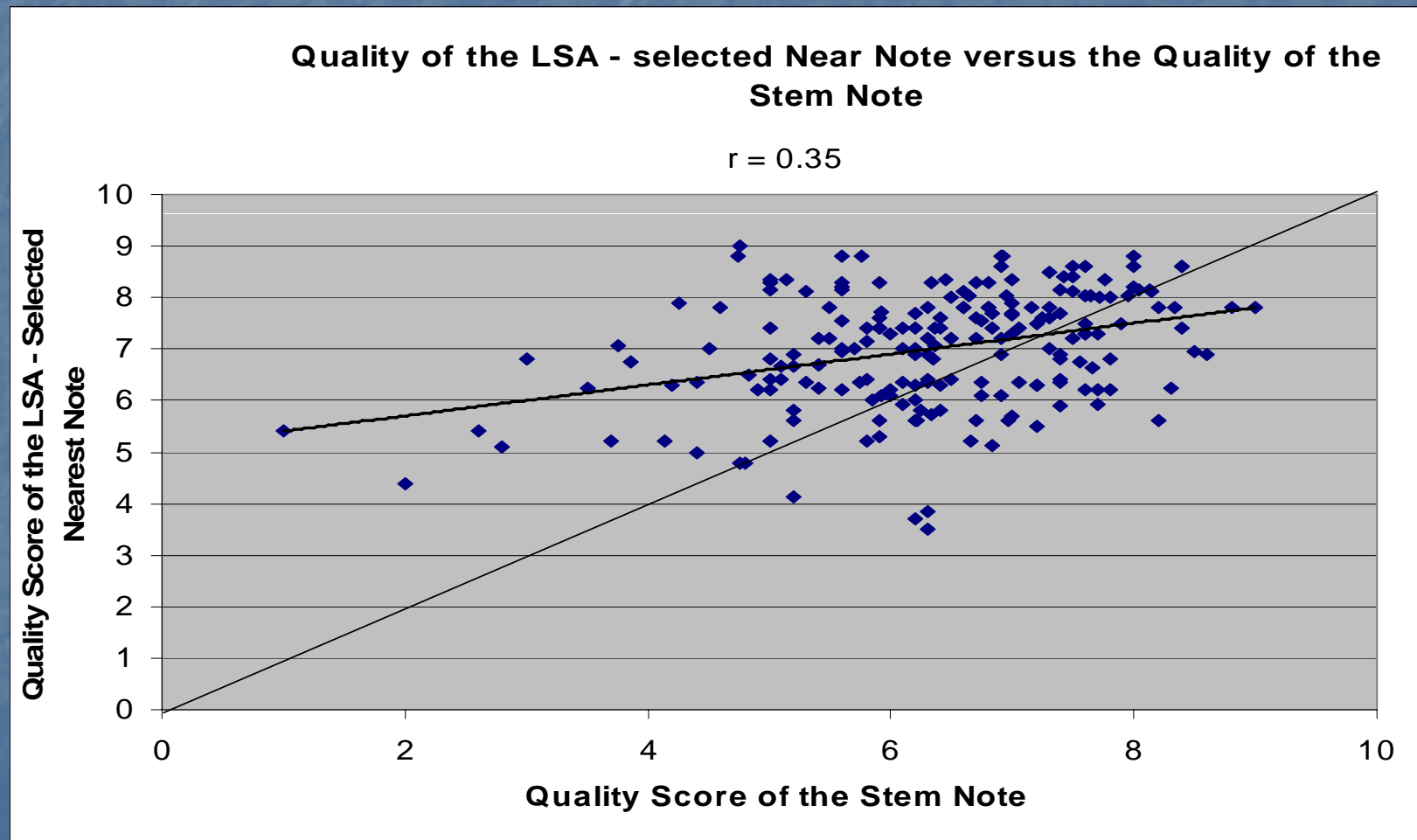
PEER PRESSURE: knowing that others will read and comment on one's solution produces more thoughtful and complete responses;

Paper vs Knowledge Post Essay Responses to TLAC1 (McLouth)

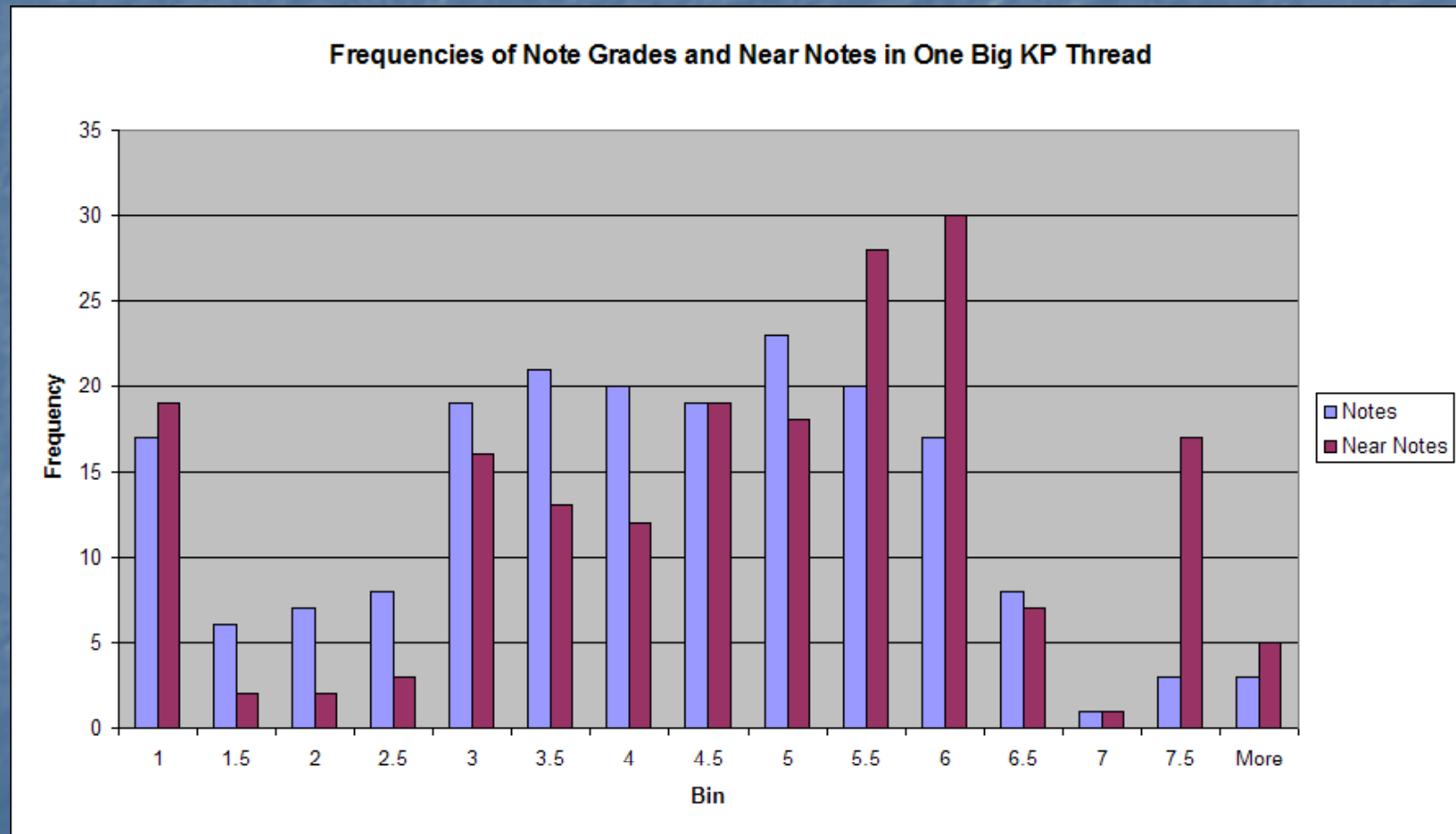


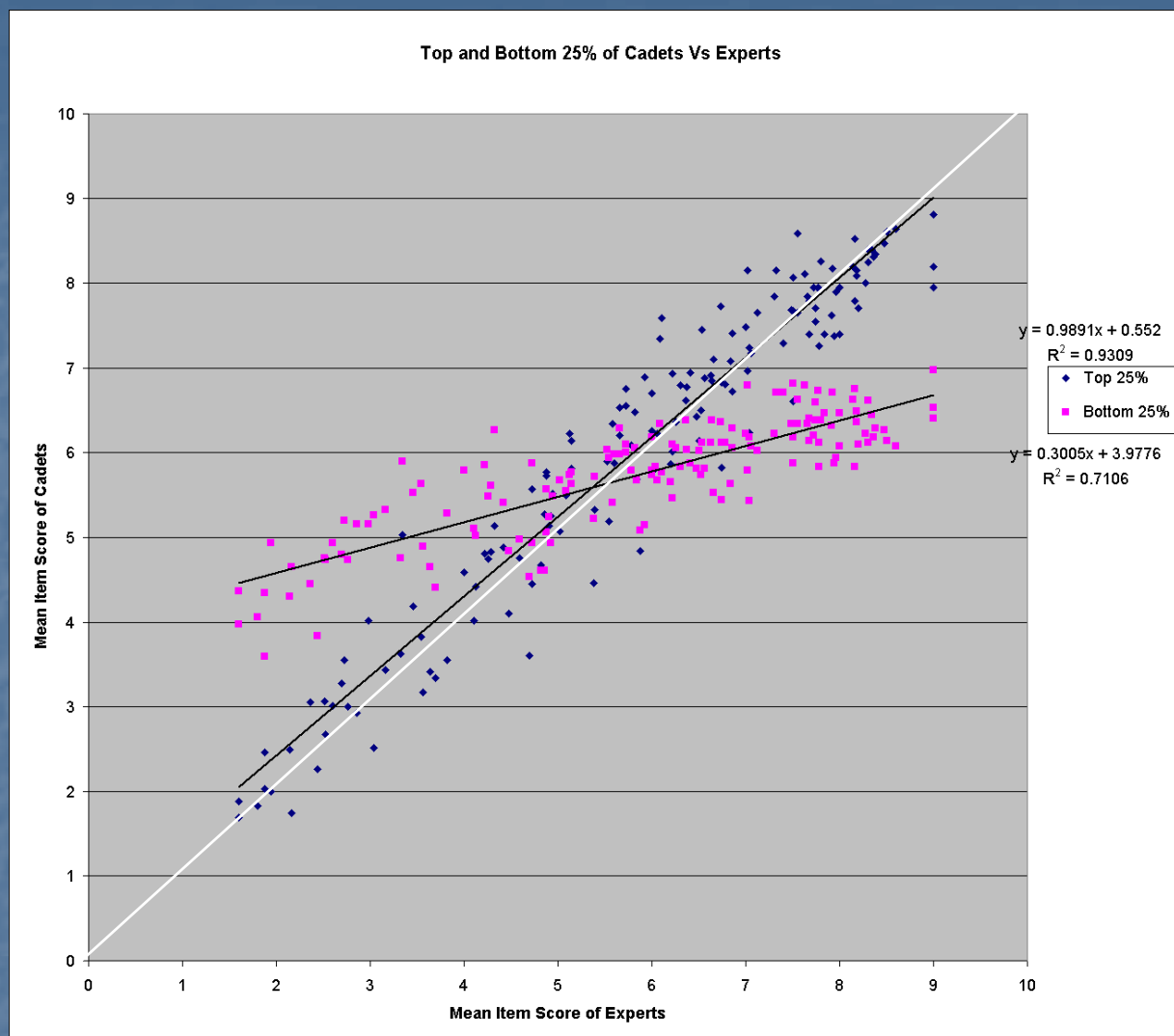
LEARNING FROM PEERS : reading and commenting, with automated assistance, on each other's notes produces a superior final solution over face-to-face.

ZPD and Goldilocks Principle Perhaps Implemented in KP by LSA



ZPD and Goldilocks Principle Perhaps Implemented in KP by LSA





The relationship between the top 25% of Cadets, the bottom 25%, and the Expert Senior Officers used to standardize the PLQ TKML, showing that the top 25% (selected using their own means) are practically indistinguishable as a group for setting the standards of the test.

Conceptualizing Consensus Based Measurement : Summary & Implications

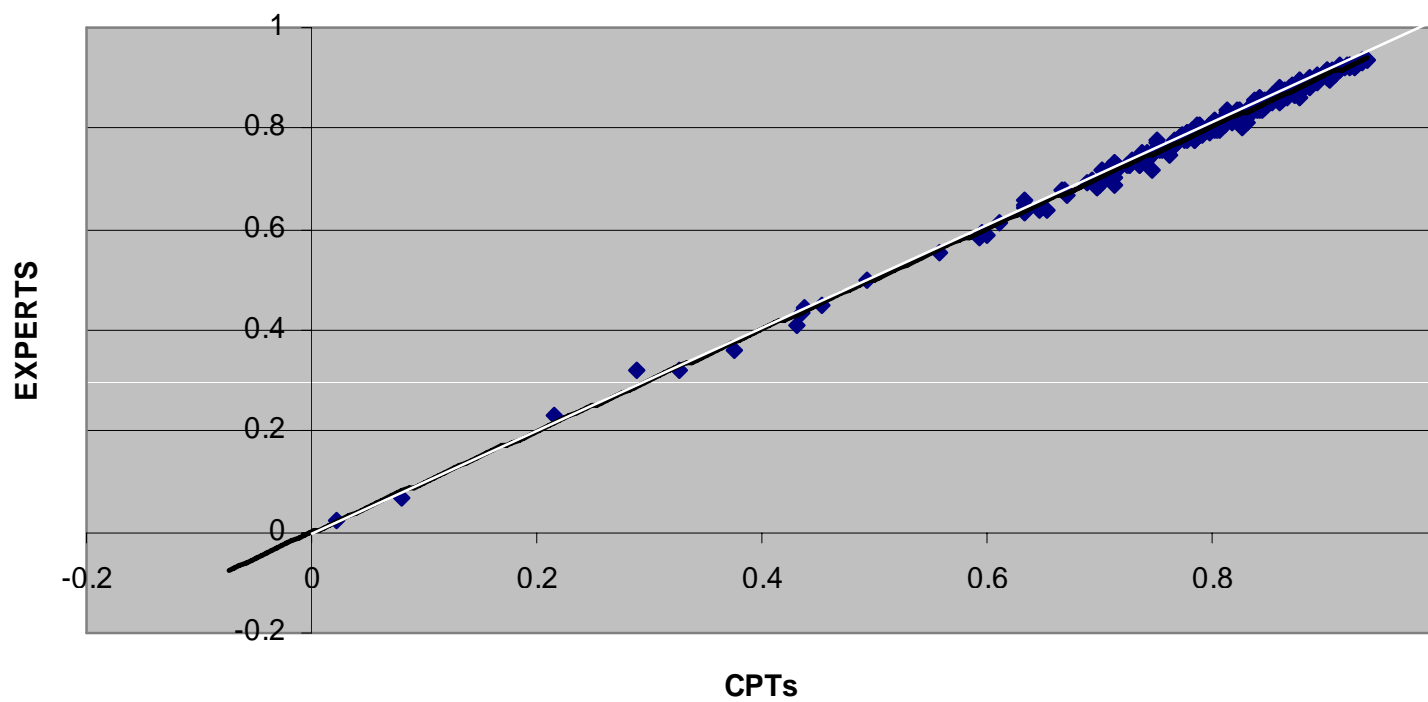


- Expert and Examinee judgments will be correlated when the space of formative experiences and tacit knowledge is similar,
- This is equivalent to the expectation that when exposure to experiences and other knowledge is similar over levels of expertise,
$$r(\text{expert, truth}) > r(\text{journeyman, truth}) > r(\text{novice, truth})$$
- Disagreement is expected when the distributions of exposure to either declarative knowledge or experiences is censored
 - Declarative Knowledge Expectations: Alcohol and crash involvement, Urban crime
 - Proscribed Experiences: Teen Smoking, Sexuality & Alcohol expectations
- Supports development of scales in domains lacking experts
- Provides economy to test development
- Explicitly invokes the concept of disagreement to understand knowledge structures



TKML Score
219 CPTs Predicted From Their OWN Overall Means Versus From Expert's Means

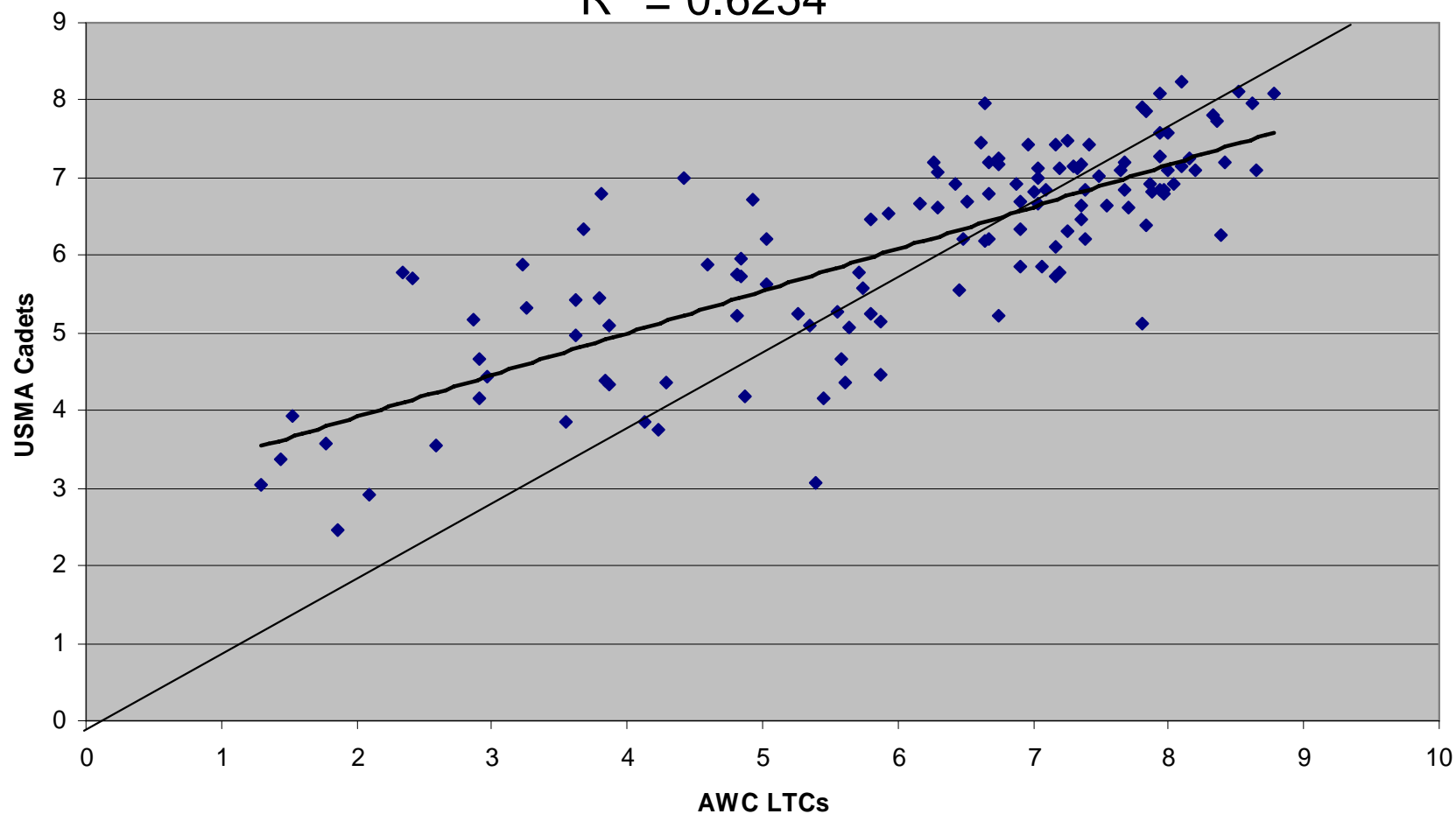
$R^2 = 0.9957$





Item Means for 33 USMA Cadets (Mn= .63) vs 33 AWC LTCs (Mn =.75) on Bn TKML

$$R^2 = 0.6254$$



Online Leader Challenges Incorporating Multimedia



Online Leader Challenges Likert Rating of Alternatives



Challenge #2 Leading in Combat

Edit your course of action selection

Pictures

Use buttons below to
view pictures from
the location of
this challenge

« Prev | Next »

Video

Click here to view
the Challenge Video



[Stream Video](#)

Your revised approach

Rest, get water, then continue on.

Select how much you agree or disagree with
taking the following courses of action:

A. Listen to the Platoon Sergeant return to FOB

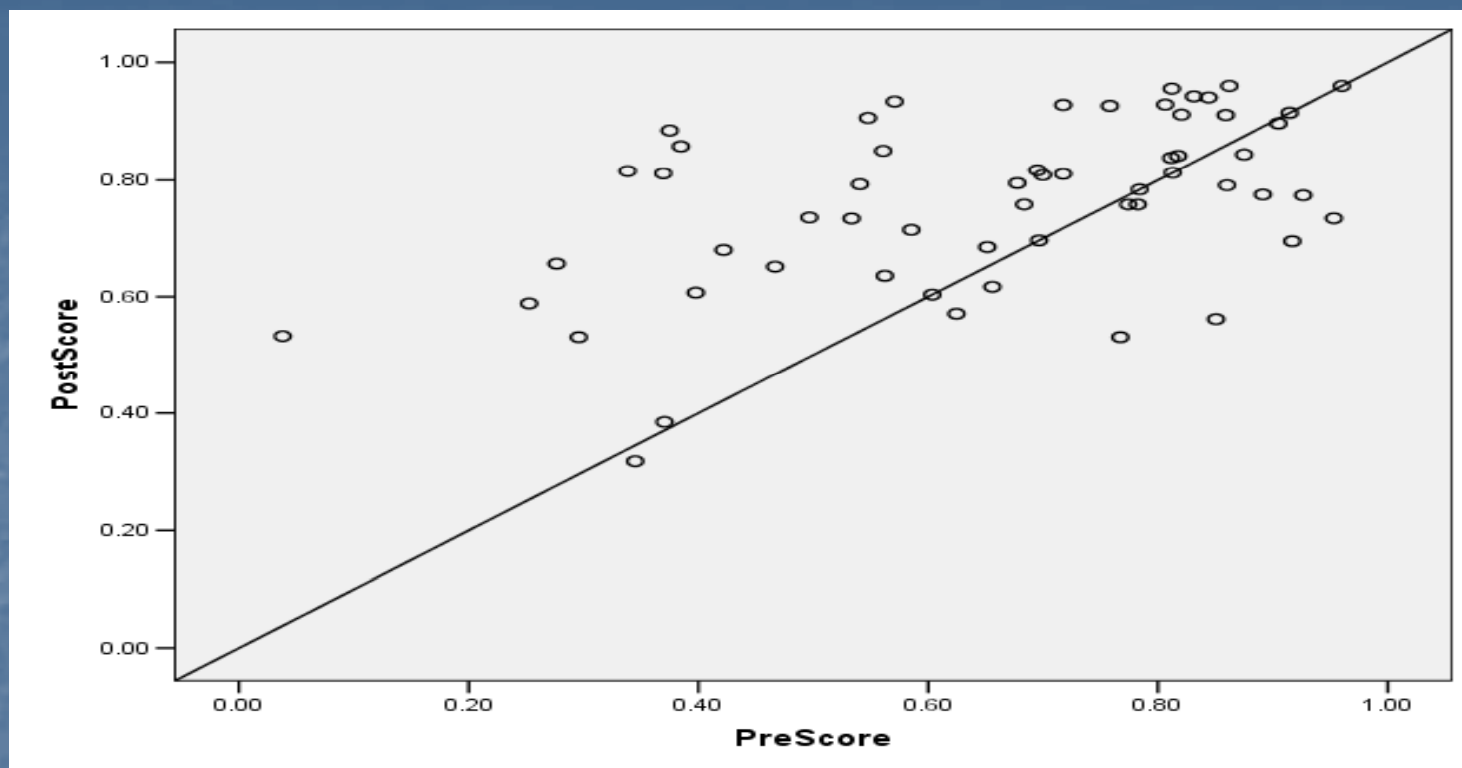
☐ strongly disagree ☐ disagree ☐ neutral ☐ agree ☐ strongly agree

B. Request that the commander clarify his task and purpose

☐ strongly disagree ☐ disagree ☐ neutral ☐ agree ☐ strongly agree

C. Take a halt to rest the platoon

☐ strongly disagree ☐ disagree ☐ neutral ☐ agree ☐ strongly agree



A comparison of pre and post-scores of USMA Cadets on Leadership SJTs scored against an expert standard.

Using the Cadet pre-test mean as the standard, the pre/post difference was still significant for LC2 and LC3 ($t(32) = 4.38, p < .001$). The important point is that a separate expert standard did not need to be derived to assess performance on this instrument: instead the scoring standard for this SJT can be derived using CBA algorithms for the Cadets themselves

1. When judgment data are collected using Likert scales, one approach to quantify individual differences is to correlate each set of respondent ratings with the scoring standard (i.e., item means).



2. These values can be computed by inverting the data matrix so that individuals correspond to columns and items to rows and then conducting a Q-factor analysis (Nunnally, 1967; Stephenson, 1935). The first set of component scores from a Principle Components Analysis (PCA) in this Q-factor analysis will correspond to the set of correlations of each individual with the scoring standard (i.e., mean ratings); we refer to these values as **component scores**.

3. Because this approach controls for respondent differences in the mean and variance of respondent ratings, these scores can be characterized as “scale-reduced.” In contrast, **the factor scores in the original PCA analysis** combine both mean and variance attributes, standardized by the amount of consensus in the scores, and so can be called “scale-loaded”.

4. These scores are broadly consistent with principles of psychophysics because they remove variance based on individual differences in the mean or “modulus” of each person’s judgments (Stevens, 1975).



Job Analysis Test (JAT) Items:

Instructions: Use all your knowledge, experience, and expertise to indicate how **frequently** each of the following tasks is performed by Soldiers at the E4 level (fully functional at skill level 10) in your occupation in a combat zone. Please use the following scale to rate how frequently most Soldiers in your occupation perform each task. Be sure to answer each question even if you have never deployed to a combat zone. Record your rating next to each item.

1. Secure the scene of a traffic accident
2. Operate a roadblock or a checkpoint
3. Supervise the establishment and operation of a dismount point

Employee Attribute Test (EAT) Items:

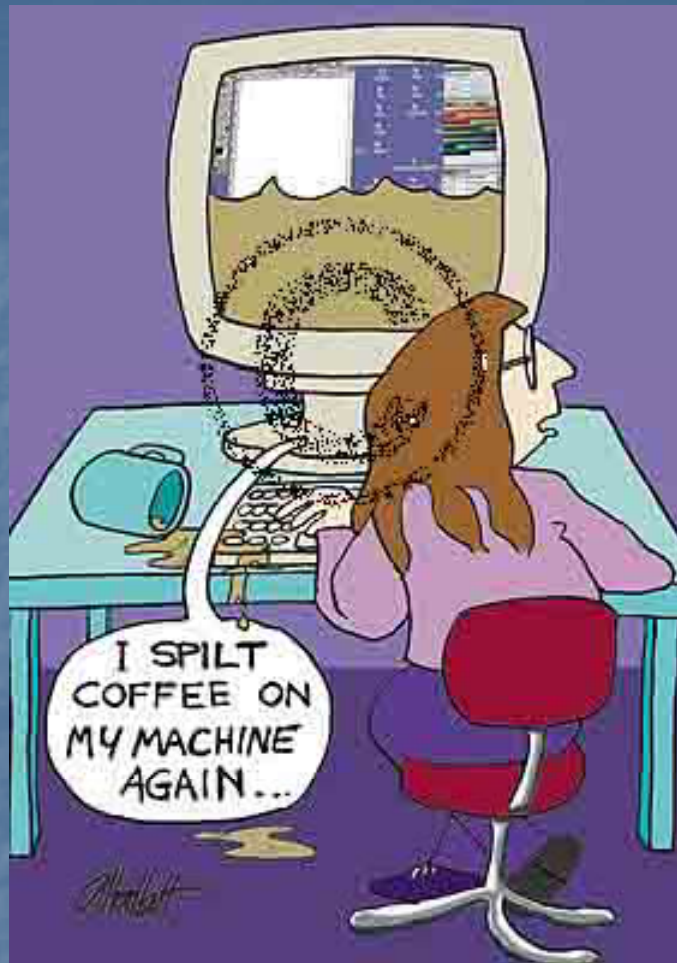
Instructions: Use all your knowledge, experience and expertise to indicate how **IMPORTANT** the Army believes each of the following characteristics is to success in your occupation at the E4/E5 level in a combat zone. Please use the following scale to rate the importance of each characteristic, and record your rating next to each item. Be sure to read the description of each characteristic and answer each question even if you have never deployed to a combat zone.

1. Conscientiousness/Dependability. The tendency to be trustworthy, reliable, and willing to accept responsibility.
2. General Cognitive Aptitude. The overall ability to understand information, identify problems & solutions, and learn.
3. Emotional Stability. Acts rationally and displays a calm mood



For both types of OJTs, the correlations between the component scores, job knowledge, and cognitive ability were statistically significant, ($p < .05$, 1-tail), as hypothesized, although the OJT component score correlations with job knowledge were significant at much more stringent levels, ($p < .001$).

The demonstration that the EAT component scores correlated with career attitudes (CA) (r [CA,EAT component] = .14, $p < .01$, $\rho = .19$) contrasts with the non-significant correlations of career attitudes with the conventional job knowledge (JK) measures, (r [CA,JK] = -.01, ns), and with AFQT, (r [CA,AFQT] = -.09, ns). These results provide the first evidence that the OJT method may assess knowledge that reflects incumbent career motivation and is acquired informally, as implied by tacit knowledge theory (Sternberg & Wagner, 1993).



"I spilt coffee on my machine"

