

# *Evaluation of Expertise in a Dynamic Decision Making Task*

James Shanteau and Clive Fullagar

Brian Friel, John Raacke, Rickey Thomas

*Kansas State University*

with David J. Weiss

*California State University, Los Angeles*

and Julia Pounds

*Federal Aviation Administration*



# *Perspectives on Expertise*

“The (expert) decision maker must continually adjust to consequences, and in doing so, deviate from the clear course laid out in advance.” (*A. Toffler, 1985*)

“A theory is only valuable if it has the ability to predict future outcomes. But no theory can predict human action.” (*M. Crichton, 1999*)



# *Purposes*

- Analysis of dynamic decision performance
- Evaluation of individual decision making
- Evaluation of team decision making
- Application of a novel measure of expertise
- Study of longitudinal development of expertise



# *Need for a Performance-Based Measure (PBM) of Expertise*

- PBM needed when no Gold Standards exist, eg, aesthetic judgments
- PBM necessary when Gold Standards have not yet been defined, eg, price for IPO's
- PBM required when there is more than one Gold Standard, eg, air traffic control



# *Previous Approaches to PBM*

- SME's: *known experts in a field*
- Between-Reliability: *cross-S consensus*
- Within-Reliability: *internal consistency*
- Discrimination: *ability to differentiate*



# *Cochran-Weiss-Shanteau (CWS)*

## *Approach to PBM*

- Based on 2 necessary conditions for expertise:  
*Discrimination and Consistency*
- Following *Cochran* (1943), a ratio is used:  
$$CWS = \text{Discrimination} / \text{Consistency}$$
- CWS provides relative, not absolute, index



# *Research Question: Longitudinal Development of Skill*

- **Question:** Can CWS be used to evaluate longitudinal skill acquisition in a complex task?
- **Comment:** Nearly all prior studies used cross-sectional (between-group) designs
- **Approach:** CWS applied to individual and team development of skill in Air Traffic Control (ATC)



# *CTEAM: Microworld Simulation of Air Traffic Control*

- **CTEAM** (Controller Teamwork Evaluation and Assessment Methodology) developed by FAA
- CTEAM “controllers” work with 1 sector in a simulated 4-sector airspace
- Controllers issue commands in a dynamic, ever-changing, real-time environment



018 AA  
019 CPIK  
020 CPIK  
021 CPIK  
022 CPIK  
023 CPIK  
024 CPIK  
025 CPIK  
026 CPIK  
027 CPIK  
028 CPIK  
029 CPIK  
030 CPIK  
031 CPIK  
032 CPIK  
033 CPIK  
034 CPIK  
035 CPIK  
036 CPIK  
037 CPIK  
038 CPIK  
039 CPIK  
040 CPIK  
041 CPIK  
042 CPIK  
043 CPIK  
044 CPIK  
045 CPIK  
046 CPIK  
047 CPIK  
048 CPIK  
049 CPIK  
050 CPIK  
051 CPIK  
052 CPIK  
053 CPIK  
054 CPIK  
055 CPIK  
056 CPIK  
057 CPIK  
058 CPIK  
059 CPIK  
060 CPIK  
061 CPIK  
062 CPIK  
063 CPIK  
064 CPIK  
065 CPIK  
066 CPIK  
067 CPIK  
068 CPIK  
069 CPIK  
070 CPIK  
071 CPIK  
072 CPIK  
073 CPIK  
074 CPIK  
075 CPIK  
076 CPIK  
077 CPIK  
078 CPIK  
079 CPIK  
080 CPIK  
081 CPIK  
082 CPIK  
083 CPIK  
084 CPIK  
085 CPIK  
086 CPIK  
087 CPIK  
088 CPIK  
089 CPIK  
090 CPIK  
091 CPIK  
092 CPIK  
093 CPIK  
094 CPIK  
095 CPIK  
096 CPIK  
097 CPIK  
098 CPIK  
099 CPIK  
100 CPIK

018 AA  
019 CPIK  
020 CPIK  
021 CPIK  
022 CPIK  
023 CPIK  
024 CPIK  
025 CPIK  
026 CPIK  
027 CPIK  
028 CPIK  
029 CPIK  
030 CPIK  
031 CPIK  
032 CPIK  
033 CPIK  
034 CPIK  
035 CPIK  
036 CPIK  
037 CPIK  
038 CPIK  
039 CPIK  
040 CPIK  
041 CPIK  
042 CPIK  
043 CPIK  
044 CPIK  
045 CPIK  
046 CPIK  
047 CPIK  
048 CPIK  
049 CPIK  
050 CPIK  
051 CPIK  
052 CPIK  
053 CPIK  
054 CPIK  
055 CPIK  
056 CPIK  
057 CPIK  
058 CPIK  
059 CPIK  
060 CPIK  
061 CPIK  
062 CPIK  
063 CPIK  
064 CPIK  
065 CPIK  
066 CPIK  
067 CPIK  
068 CPIK  
069 CPIK  
070 CPIK  
071 CPIK  
072 CPIK  
073 CPIK  
074 CPIK  
075 CPIK  
076 CPIK  
077 CPIK  
078 CPIK  
079 CPIK  
080 CPIK  
081 CPIK  
082 CPIK  
083 CPIK  
084 CPIK  
085 CPIK  
086 CPIK  
087 CPIK  
088 CPIK  
089 CPIK  
090 CPIK  
091 CPIK  
092 CPIK  
093 CPIK  
094 CPIK  
095 CPIK  
096 CPIK  
097 CPIK  
098 CPIK  
099 CPIK  
100 CPIK

021 AA  
S1 ICP9  
015 AA  
S1 PCIK

023 AA  
S1 IICd

009 00  
P4 I2

006 00  
S3 PCIK  
035 00  
S1 IICd

028 00  
S4 IICd  
029 00  
S1 ICP9

039 00  
01 I2  
017 00  
027 00  
001 IFCB  
047 00  
042 K  
043 00  
01 I2  
020 00  
S4 IICd  
012 00  
044 CLIC  
025 00  
001 IFCB

012 00  
044 CLIC

025 00  
001 IFCB

045 BB  
001 e

000 033 BB  
P4 ICP9  
044 BB  
P4 ICP9

037 BB  
044 PG

024 BB  
P4 CLK

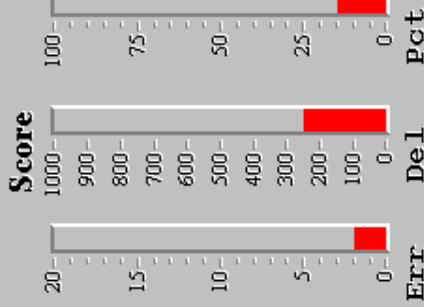
011 CC  
P3 P8

034 CC  
P4 P8  
042 CC  
S1 G  
038 CC  
S1 I

046 CC  
P4 h

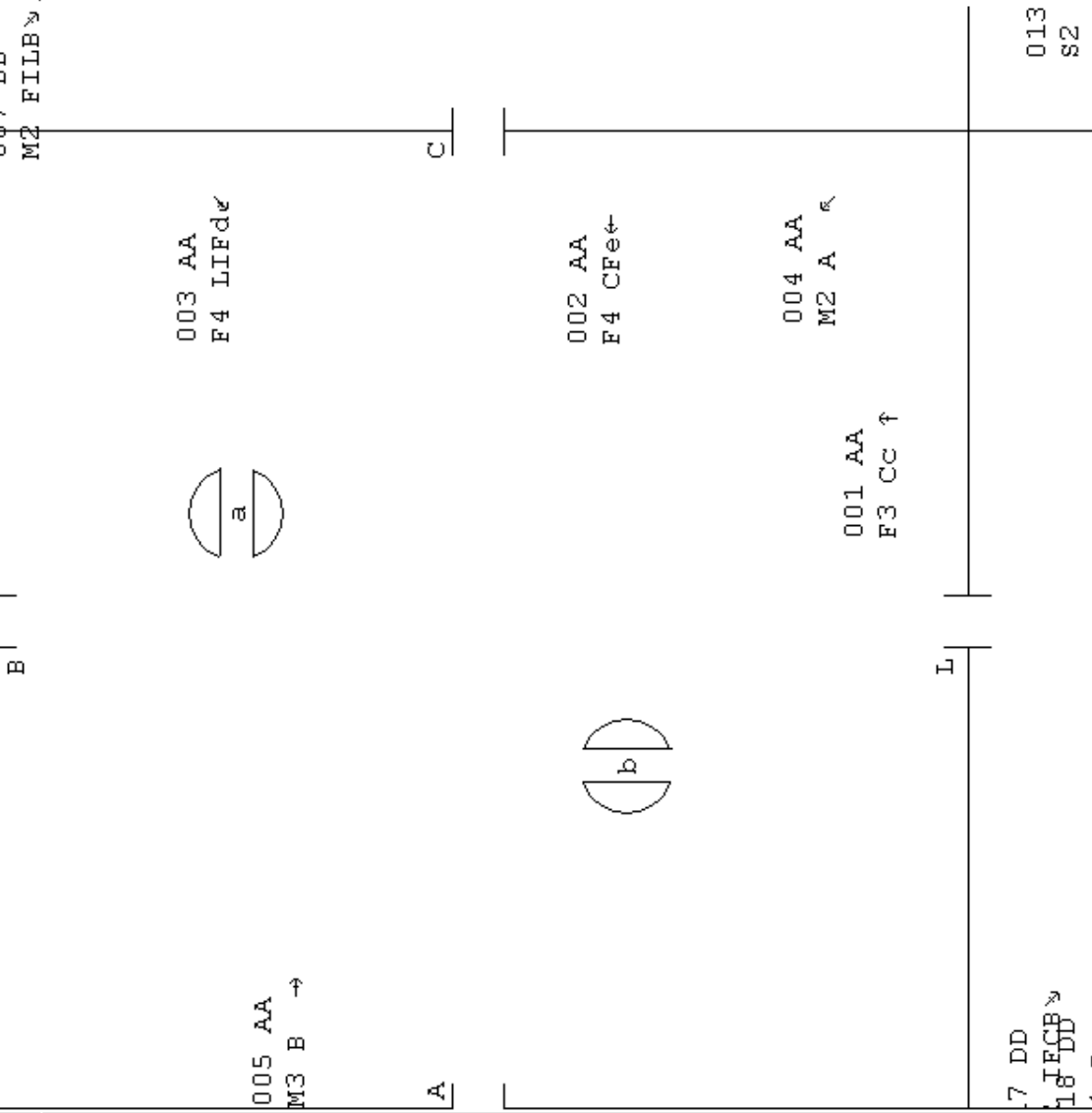
**SECTOR INFO**

Landing Time  
 a → 1620  
 b ↑ 0000000



**Pilot Readback**

Roger, Level 3  
 Sector Message  
 From BB: Request  
 H/O, Gate C, AC 013



**COMMAND**

REPEAT

Heading  
 7 0 1 2  
 6 5 4 3

Speed  
 F  
 M  
 S

Level  
 4  
 3  
 2  
 1

Handoff  
 To  
 AA BB  
 DD CC

5 Miles

**INTER-SECTOR MESSAGE**

To BB Request Handoff Control  
 DD CC Accept Traffic None  
 Reject Delay Restrict

Where  
 AA C BB  
 L F  
 DD I CC

With Who  
 AA BB CC  
 001 006 011  
 002 007 012  
 003 008 013

With Restrictions  
 Heading Speed Level  
 7 0 1 4  
 6 5 4 3  
 5 4 3 2  
 1

SEND  
 CANCEL

# *Study I*

- 12 undergraduate students at KSU spent 3 months with CTEAM
- Individual performance evaluated in single-sector version of CTEAM
- CWS scores computed individually

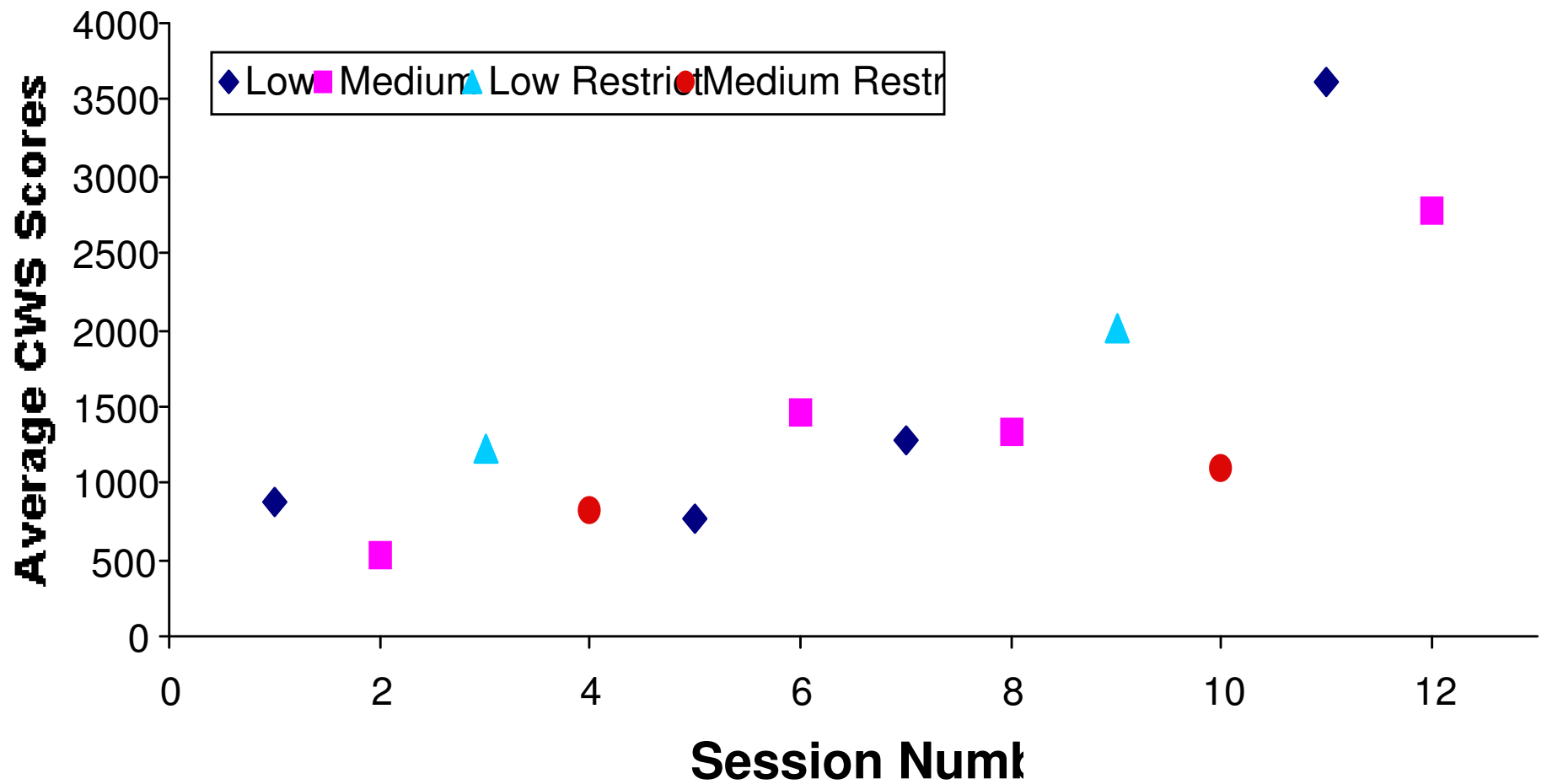


# *Methodology*

- Three independent variables:
  - *Density*: Low(12 aircraft) and Medium(24)
  - *Restricted airspace*: Yes or No
  - *Sessions*:12 sessions @ 2 hours/session
- *Design*: Each scenario repeated 3 times/sess
- Three dependent variables:
  - Number of separation errors and crashes
  - Number of control actions
  - Time to destination = Time Through Sector



# CWS Scores for Longitudinal

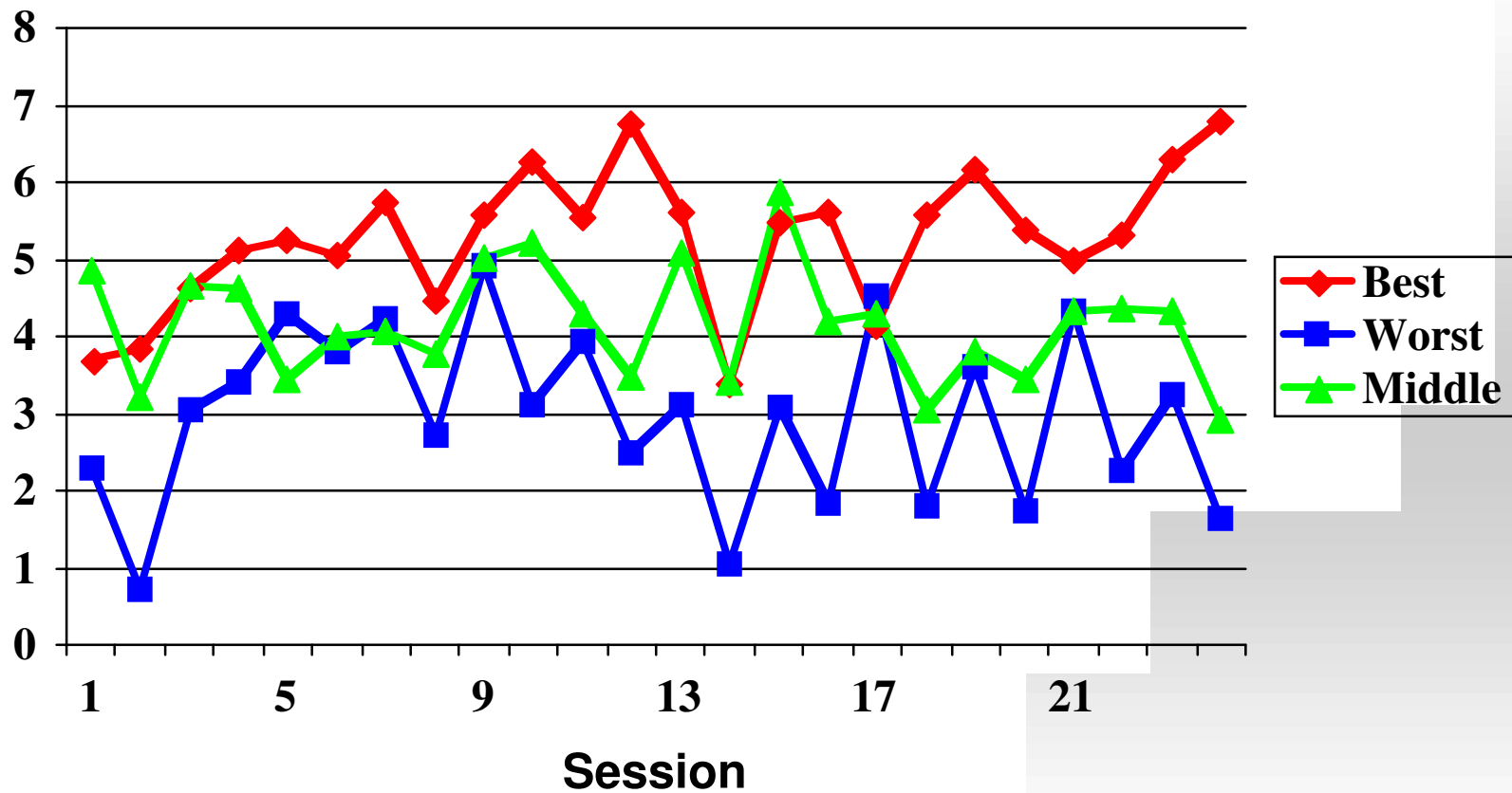


## *Study II*

- Longitudinal Study of 3 Teams of 4 participants
- 3 repetitions of the same scenario / session
- 24 sessions over 8 weeks



# *Team Performance Scores Across Sessions*

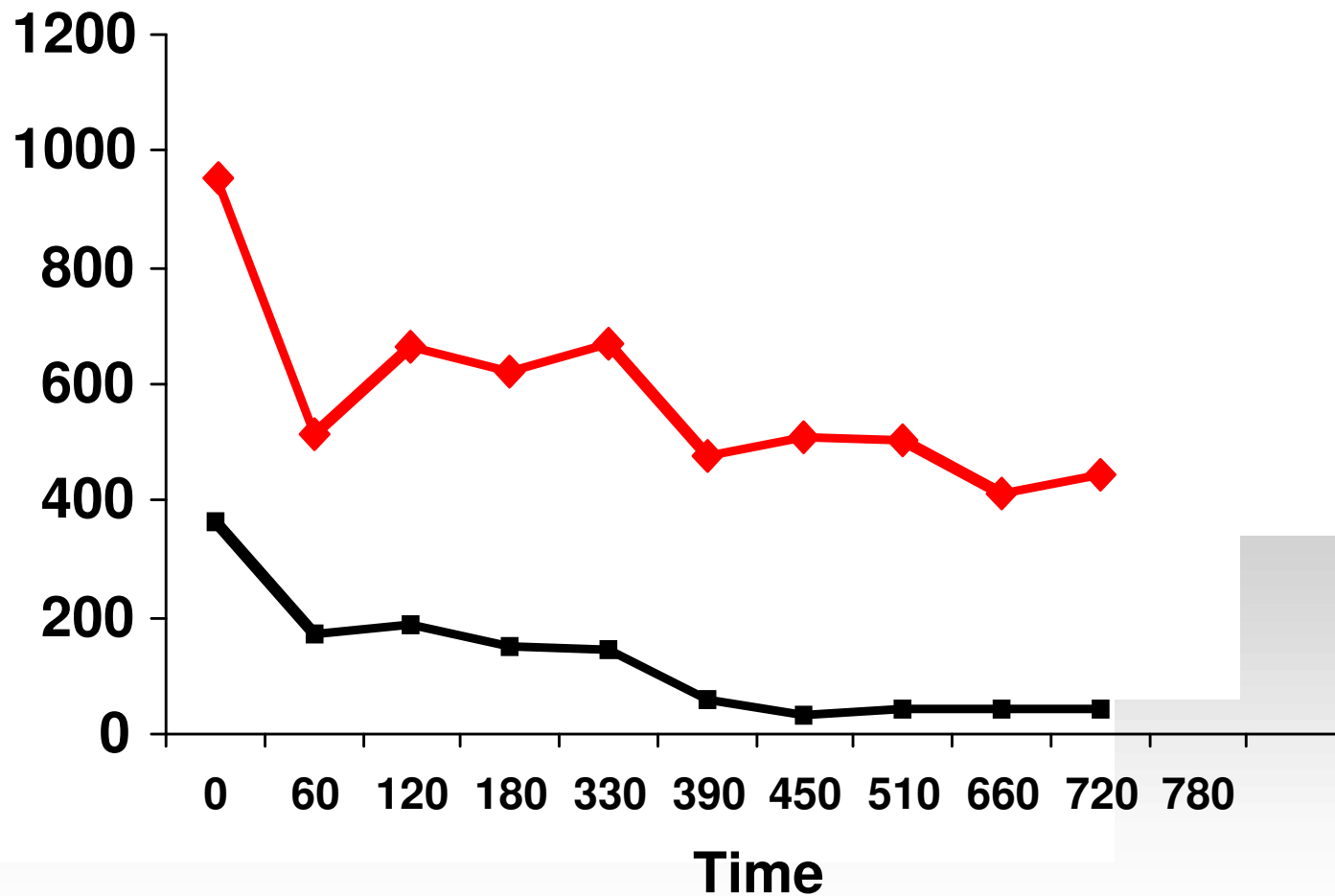


# *Study III*

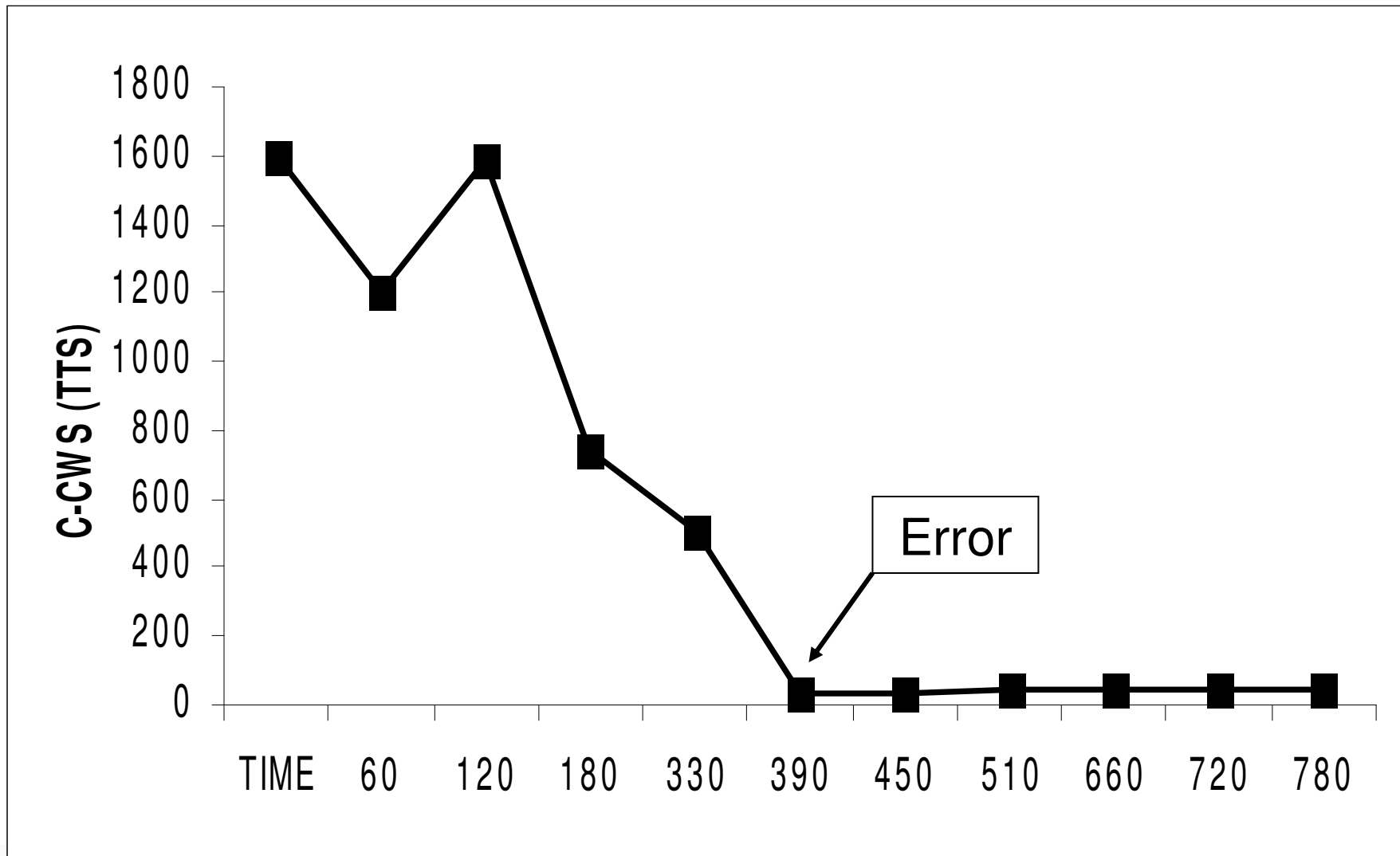
- Use of Continuous-CWS scores = C-CWS
- 12 participants run for 24 sessions in 8 weeks
- C-CWS evaluated between/within sessions



# *Longitudinal Analysis*



# *Dip in C-CWS Precedes an Error*



# *Conclusions*

- CWS successful in evaluating longitudinal development in real-time, dynamic tasks
- CWS applied both to individuals and to teams
- CWS has been successful in various domains, eg, ATC, agriculture, medicine, auditing
- CWS superior to other measures of expert performance, eg, consensus or SME ratings
- CWS works when there is no “gold standard”



# *Extensions*

- CWS can be applied to selection, ie, good early performance predicts good later performance
- CWS can be applied to evaluate training effectiveness and to identify areas where more training is needed
- CWS can be applied to predictive assessment, ie, to look ahead to “predict errors”



# *Caveats*

- CWS is situation/task specific
- CWS scores cannot be compared across tasks
- CWS cannot be applied to one-time tasks



# *For More Information*

- There are downloadable papers and free programs on both CWS and C-CWS
- There is a Workbook (pdf format) on how to use and interpret CWS
- All of these can be obtained at our website:  
<http://www.ksu.edu/psych/cws/>



# Quotes

*An expert is someone who knows some of the worst mistakes that can be made in his subject and how to avoid them (Heisenberg)*

*An expert is somebody who is more than 50 miles from home, has no responsibility for implementing advice he gives, & shows slides (Meese)*

