

# A Beautiful Mind is A Terrible Thing to Lose

*James Shanteau (KSU)*

*David Weiss (CSU-LA)*

*Rickey Thomas, Brian Friel,*

*John Raacke (KSU)*

*Julia Pounds (CAMI, FAA)*



# What are the Limits of Expertise?

After winning the 1996 Nobel Prize in physics, Douglas Osheroff was asked for advice on K-12 science education by a legislative aide.

*“I’m not an expert in that,”* said Osheroff.

*“Certainly you must be,”* replied the aide.

*“You’re a Nobel laureate.”*

This conversation illustrates a commonly-held view that expertise is like a personality trait, ie, someone either is, or is not, an expert.



# Broad and Narrow Generalizability of Expertise

When faced with unfamiliar tasks, experts should be able to adapt, eg, good classroom teachers should be able to teach effectively on the Internet

==> *Broad Generalizability*

Experts should be able to extend their skills to closely related tasks, eg, good algebra teachers should be able to teach geometry

==> *Narrow Generalizability*



# Tests of Predictions

Using reanalyses of prior studies of experts, we will test these two predictions:

- n *Prediction# 1:* Experts should be able to adapt to unfamiliar tasks, ie, their skills should be broadly applicable.
- n *Prediction #2:* Experts should be able to perform closely related tasks, ie, their skills should extend to similar problems.



# Cochran-Weiss-Shanteau (CWS) Index of Expert Performance

Based on 2 necessary conditions for expertise:

*Discrimination and Consistency*

$CWS = \text{Discrimination} / \text{Inconsistency}$

Higher CWS scores implies better performance

Can use any summary statistic, eg, variance

CWS provides relative, not absolute, index

Can only be used within, not across, domains



# Example Calculation of CWS

Doctor #12:	A	B	C	D	E
Repl # 1:	100	2	95	98	28
Repl # 2:	98	15	91	98	54

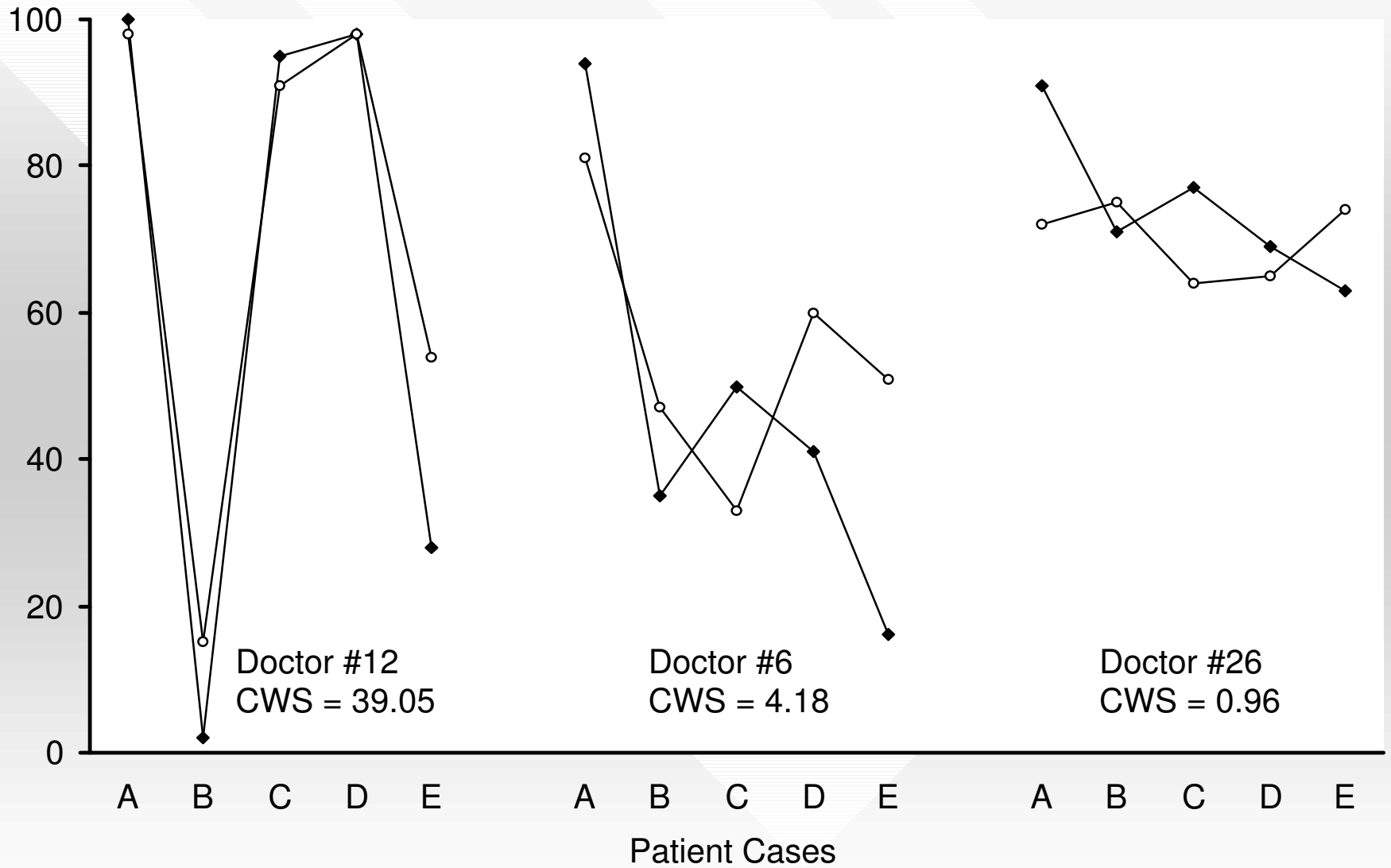
(from *Skånér, et al, 1998*)

$$\text{Discrimination} = ? \quad (M_i - GM)^2 / n - 1 = 3377.60$$

$$\text{Consistency} = ? \quad (d_j^2) / n = 86.50$$

$$\text{CWS} = 3377.60 / 86.50 = \mathbf{39.05}$$





# Prediction #1: Broad Generality of Skills to Other Tasks in Domain

Thomas, Willems, Shanteau, Raacke, & Friel (2001) reanalyzed an FAA study of 12 level-5 air traffic controllers (ATCs) conducted by WJH Technical Center.

Airspace complexity and conflict type varied in simulated ATC scenarios.

“Objective Performance” = Separation Errors.

Subject Matter Experts (SMEs) rated ATCs

Correlation between SMEs & errors =  $-.14^{ns}$

Correlation between CWS & errors =  $-.47^*$



# Why Aren't SMEs Expert at Rating Other Controllers?

CWS analyses of ATC performance were more sensitive to errors than SME ratings. Why?

SMEs were asked to rate controllers on a wide variety of scales ranging from Efficiency and Conflict Resolution to Knowledge and SA.

Why should SMEs be expert at rating ATCs?

*Most great players don't know how to coach.*

*Prediction #1 not supported:* SMEs were not able to extend their skills to a different task.



## *Prediction #2: Narrow Generality of Skills to a Closely Related Task*

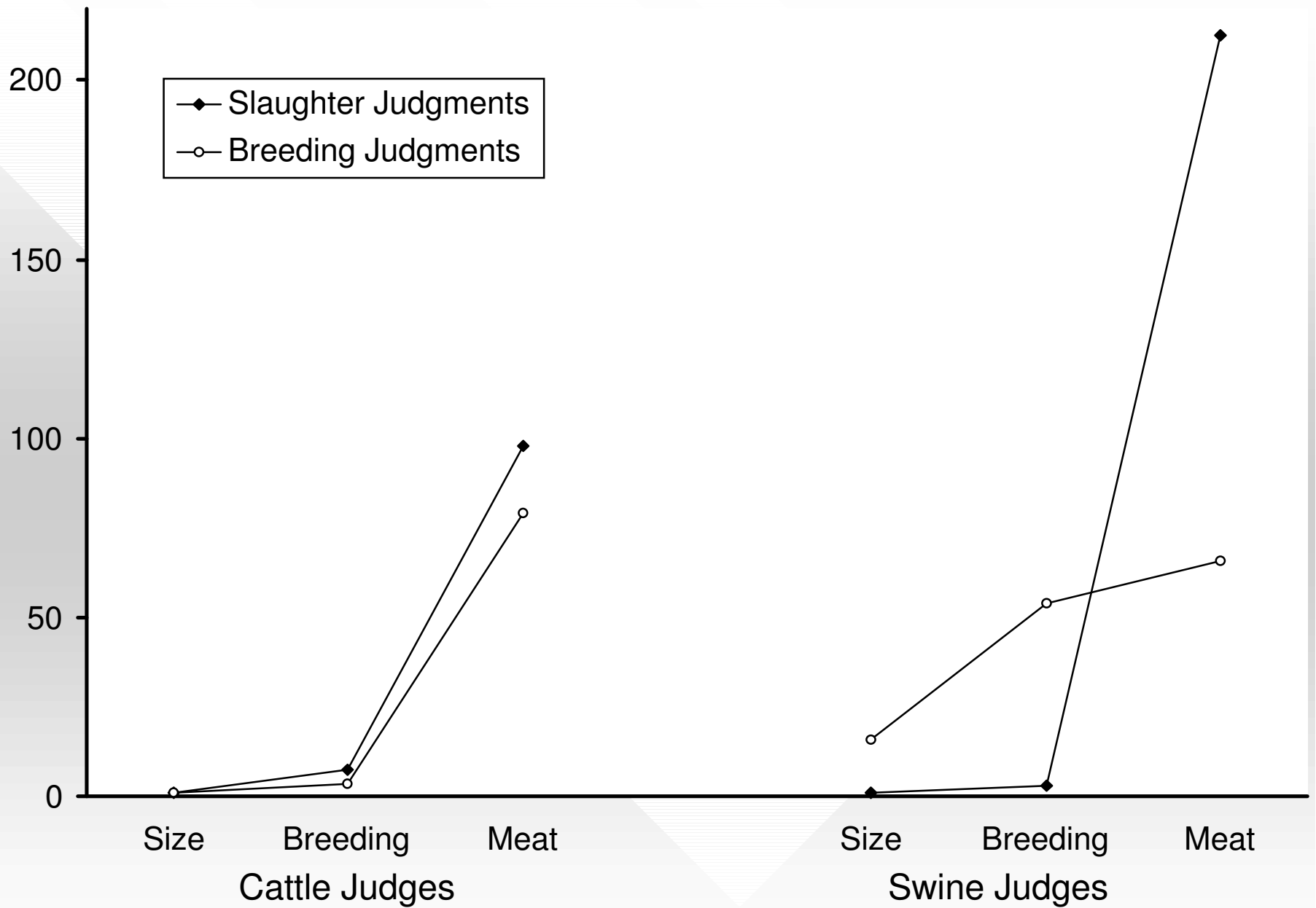
Phelps asked 4 professional livestock judges to evaluate 27 drawings of gilts (female pigs) for slaughter quality and for breeding quality.

Two of the judges were experts in swine and two were experts in cattle.

Drawings varied systematically as a function of size, breeding, and meat.

Drawings were shown 3 times to each judge.





# Why the Big Difference?

Cattle judges were similar to swine judges when judging slaughter quality. But, they were different when judging breeding.

Although cattle judges are familiar with swine breeding, they lack day-to-day experience.

This highlights importance of task specificity.

*Prediction #2 not supported:* Livestock judging skills do not generalize from cattle to swine.

*Expertise is spectacularly narrow (Chase).*



# Other Findings

Weiss (2000) examined ratings by ice skating judges of artistic impression and technique. Although supposedly independent, the ratings were highly correlated ( $r = .95$ ). Basically, the judges made one evaluation – *twice!*

Abdolmohammadi & Shanteau (1999) asked auditors to rate characteristics of “experts” with different specialties, eg, manufacturing vs non-profits. Few differences were found – the same characteristics were rated high for all specialties, eg, communication skills.



# Specificity vs Generality

Counter to common belief, expertise is tightly tied to very narrow tasks.

This should not be surprising – experts specialize in very specific tasks.

Would you want a heart surgeon to spend time learning about psychiatry?

*Let them be the best they can be – in their specific field!*



# What's Missing?

The generality assumed in expertise is not supported by evidence. There does not appear to be any generality across tasks.

What is needed is a taxonomy of tasks.

*We have no choice but to develop a taxonomy of intellectual tasks themselves. Only with the aid of such a taxonomy can we think with reasonable sophistication about how to identify...the myriad types of experts (Edwards, 1983).*



## 2 Partial Attempts at a Taxonomy

Shanteau (1992, 2000) proposed a taxonomy of expert domains based on degree of decision support available:

Aided > Competent > Restricted > Random

Weiss (2001) developed a hierarchical taxonomy: All tasks require Evaluation plus...

Forecasting = Evaluation + Predication

Instruction = Evaluation + Communication

Performance = Evaluation + Execution

However, neither approach is specific enough.



# Is Anything Generalizable?

Both broad and narrow tests failed to support generalizability of expertise. Neither skills nor findings generalize.

The one element that did generalize across studies was the usefulness of CWS as a metric of expertise.

We may not know what experts have in common (if anything), but we do know that CWS is a useful measure in all cases.



# Reasons for Belief in Generality

- n Desire of the public to get answers to difficult questions from someone – anyone. We want experts to be all-knowing.
- n Society values labels of expertise, eg, Nobel.
- n Desire of experts to “push the envelope” of their skills. Experts want to test their limits to find out how good they really are.
- n Some experts actually believe that their expertise in one area gives them expertise in other areas, eg, arrogant surgeons.



# Why Do We Expect Generalization?

## Two Psychological Constructs

- n Spearman distinguished general ability (*little g*) from task-specific ability (*little s*). Parallel arguments applied to expertise. Hence, *little e* (for generalizable expertise) assumed to be large & widely applicable.
- n *Generalization gradient* observed in learning, perception, etc. Initially, the tendency to generalize is large ==> wide gradient. Further discrimination training needed to reduce size of the gradient.



# Conclusions

- n Experts are “experts” because they are good at something highly specific.
- n We should not assume there is anything in common across those labeled as “expert.”
- n Expertise is not a property of the person. Rather it is a skill shown in specific tasks.
- n We need to have a taxonomy of tasks performed by experts before generalizing.
- n CWS is useful in a wide variety of tasks. It allows for direct assessment of expertise.



# Closing Comments on Experts

*An expert is one who knows everything about something and nothing about anything else (Ambrose Bierce, 1911)*

*An expert is a man who has made all the mistakes which can be made in a very narrow field (Niels Bohr, 1962)*

*The goal is not to evaluate experts. Instead, the goal is to evaluate performance of experts on specific tasks. (Julia Pounds, 2000)*

