Improving Judgment Performance by Examining the Relationship Between Task Properties and Cognitive Mode

Dr. Robert Holcomb

Dr. Leonard Adelman

20 June 2012

Introduction

- Purpose of this research is to illuminate the degree to which task structure influences cognitive mode and task performance in judgment tasks.
- Underlying basis for research is Hammond's Cognitive Continuum Theory (CCT)
  - Cognition runs in a continuum from intuitive to analytical
  - Task properties run in a parallel continuum, and when task properties match cognitive mode, achievement is improved
- Previous research in CCT has seen mixed empirical support.
  - Hammond et al. experiment in 1987; showed some support
  - Dunwoody et al. experiment in 2000; showed mixed results
- Previous experiments used different metrics and had different, and unexpected, results
- We developed a new metric that did demonstrate empirical support for CCT

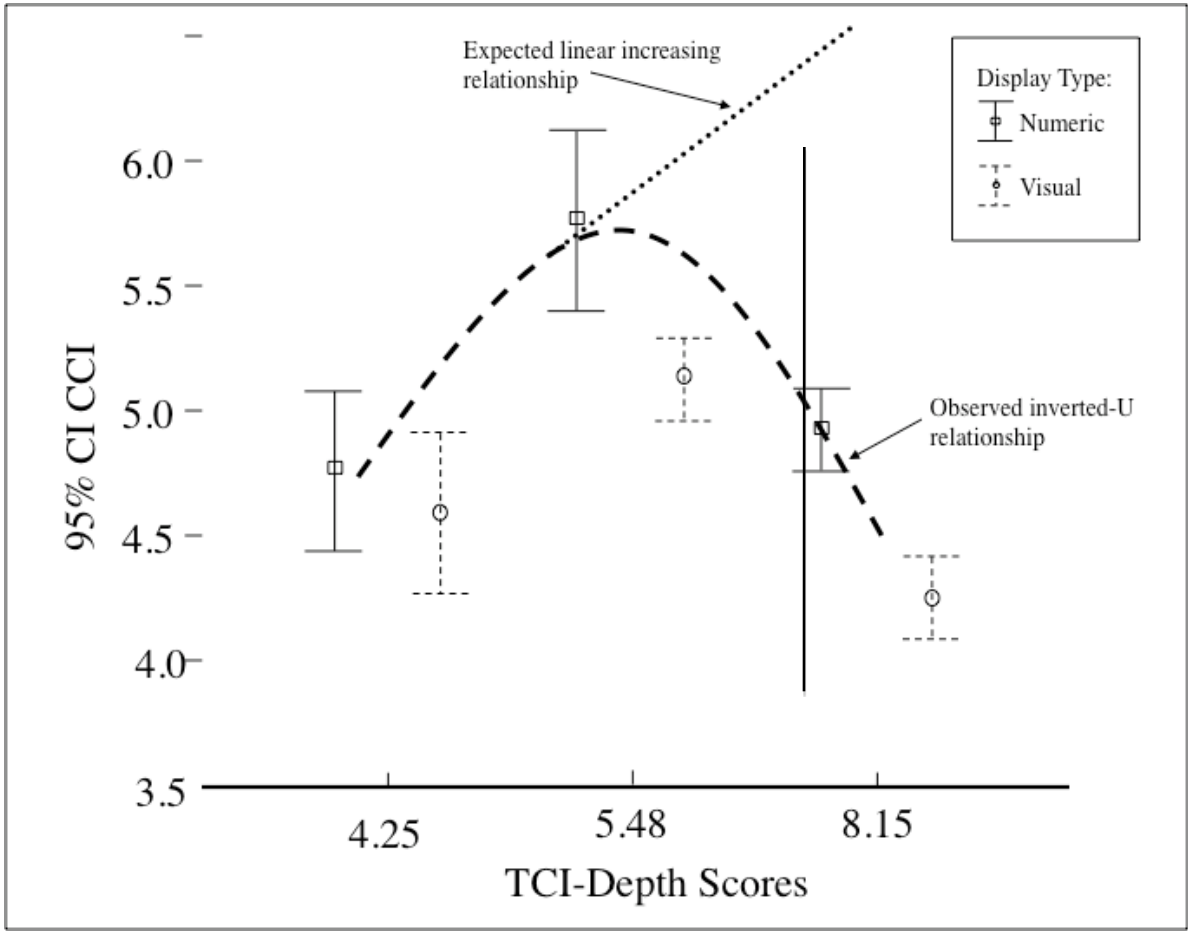
Why Is This Important to C^2?

- An empirically-supportable theory relating task structure to human cognition can result in higher task achievement for judgment tasks
- Matching task properties with the corresponding cognitive mode (e.g., intuitive or analytical cognition) can improve efficiency by predicting improved achievement
- These benefits can improve the design of command and control systems by incorporating elements of cognitive systems engineering

Hammond et al. Experiment (1987)

- Tested relationship between task properties and cognitive mode for highway engineers
- Tested three *task surface* characteristics (film strips, bar graphs and formulas) against three *task depth* characteristics (judging highway aesthetics, safety and capacity)
- Five hypotheses, of which three are relevant to this discussion:
  - H_1: Surface and depth task properties induced corresponding cognitive mode
    - Supported in most cases but not all
  - H_2: Intuitive cognition could outperform analytical cognition
    - Supported, demonstrating that one cognitive mode is not always best
  - H_3: Knowledge of the congruence between surface and depth task characteristics would be necessary and sufficient to predict achievement
    - Not supported in their results

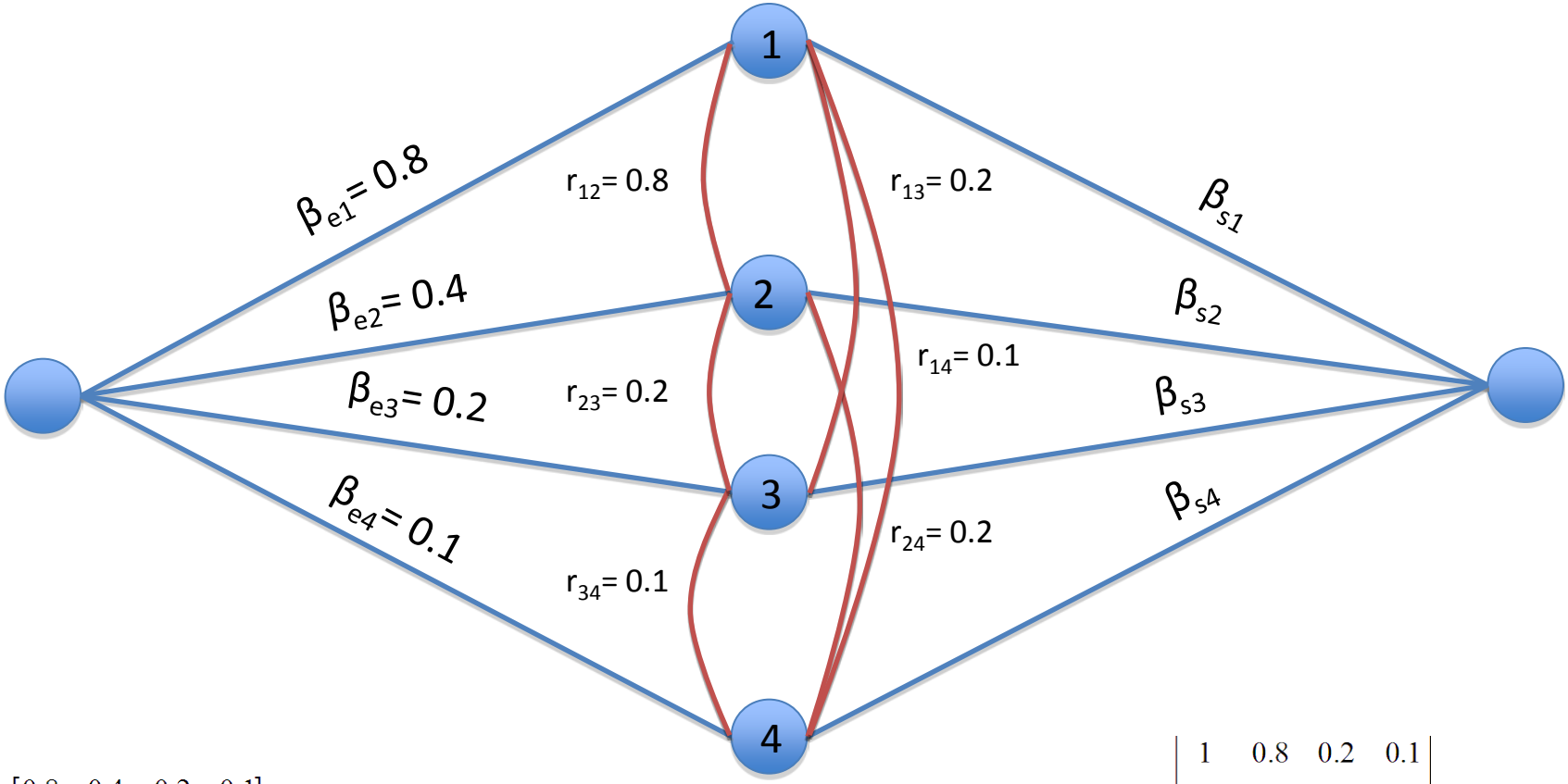
Dunwoody et al. Experiment (2000)



Our Issues with Previous Experiments

- Construction of TCI and CCI indices differed between Hammond et al. and Dunwoody et al.
  - Hammond et al. describes eleven relevant task properties, but only used eight in their experiment
  - Dunwoody et al. used four of the eight in their task index
  - Similar issue with both CCI indices of cognitive mode
  - Selection of task properties to include in their respective indices seems arbitrary in both cases
- Hammond et al. results are relatively weak, preserving order but not precise location on the continuums
- Dunwoody et al. results were unexpected, with cognitive mode on analytical task very similar to cognitive mode on intuitive task

An Example



$$\beta e = [0.8 \quad 0.4 \quad 0.2 \quad 0.1]$$

$$R = \begin{bmatrix} 1 & 0.8 & 0.2 & 0.1 \\ 0.8 & 1 & 0.2 & 0.2 \\ 0.2 & 0.2 & 1 & 0.1 \\ 0.1 & 0.2 & 0.1 & 1 \end{bmatrix}$$

Research Question

Does an index based on a formulation of vicarious mediation and vicarious functioning outperform Dunwoody et al.'s index in an empirical demonstration of Cognitive Continuum Theory?

VMI Index Construction

- Before an experiment could be devised, some preliminary work had to be accomplished:
  - what is a good operational definition of variability in matrix product?
  - what values of ecological validity weights (β_e) are operationally meaningful?
  - How should one rank order a set of matrices of non-uniform **R**?
- We chose the *mean deviation* of the matrix product of β_e**R**

$$Mean\ deviation = \frac{\sum \left| X_i - \bar{X} \right|}{n}$$

  - a more robust estimator of variability than range alone
- We chose common heuristics seen in judgment research as choices for ecological validity (β_e) values
  - Take the Best heuristic: $\beta_{e1} > \Sigma\ (\beta_{e2}+\beta_{e3}+\beta_{e4})$
  - Tally heuristic: $\beta_{e1} = \beta_{e2} = \beta_{e3} = \beta_{e4}$
- We chose (1 – det[**R**]) as a scalar way to rank a set of matrices

Task Packages

Variable	Task Package 1 (Intuition)	Task Package 2 (Quasi-rational)	Task Package 3 (Analytical)
Ecological Validities	Tally (0.25 for all)	Take the Best (0.8,0.4,0.2,0.1)	Take the Best (-0.8,0.4,0.2,0.1)
Cue Intercorrelations	**R** uniform (0.1)	**R** non-uniform (0.8,0.2,0.1)	**R** non-uniform (-0.7,0.2,0.1)
VMI score	0	0.385	0.625
average r_{ij} (used in TCI_D)	0.1	0.267	-0.060
Std Dev of βe (used in TCI_D)	0	0.310	0.532
Task Predictability	R^2_e=0.55	R^2_e=0.59	R^2_e=0.94
Overall TCI_D score	0.51	0.56	0.78

Methodology (1)

- We performed an experiment using both the VMI index and Dunwoody et al.'s TCI index and compared the results
- The experiment used a double system lens model design with four cues
- We controlled environment side of the lens through task properties as Dunwoody et al. did.
- We created three task packages (inducing intuition, quasi-rationality and analysis) presented to participants
- Participants made judgments based on the four cues
- We used a within-subjects design (as in Hammond et al.) vice between-subjects design (as in Dunwoody et al.)
  - cognitive shift across tasks better observed with a within-subjects design
  - Analysis done both within-subjects and between-subjects

Methodology (2)

- Participants were experienced teachers performing a student placement task on hypothetical student profiles (60 student profiles for each of the three tasks)
  - Demographics: 43 females, 9 males, average 17.3 years experience teaching
- Student profile information was created by varying the task properties of β_e and [**R**] in the three cases described above
  - We used examples drawn from the Scales for Rating the Behavioral Characteristics for Superior Students, familiar to the subjects
  - Tasks were labeled as judgments on students creativity, academic performance and learning potential
  - Tasks were presented to teachers as either bar graphs or tables of values
- A pilot test was done to be sure tasks were familiar to the subjects; no issues arose in pilot test
- Participants volunteered for the study and were awarded continuing education points based on their participation
- Participants performed the judgments, recorded their times and demographic information, then returned the packages via mail

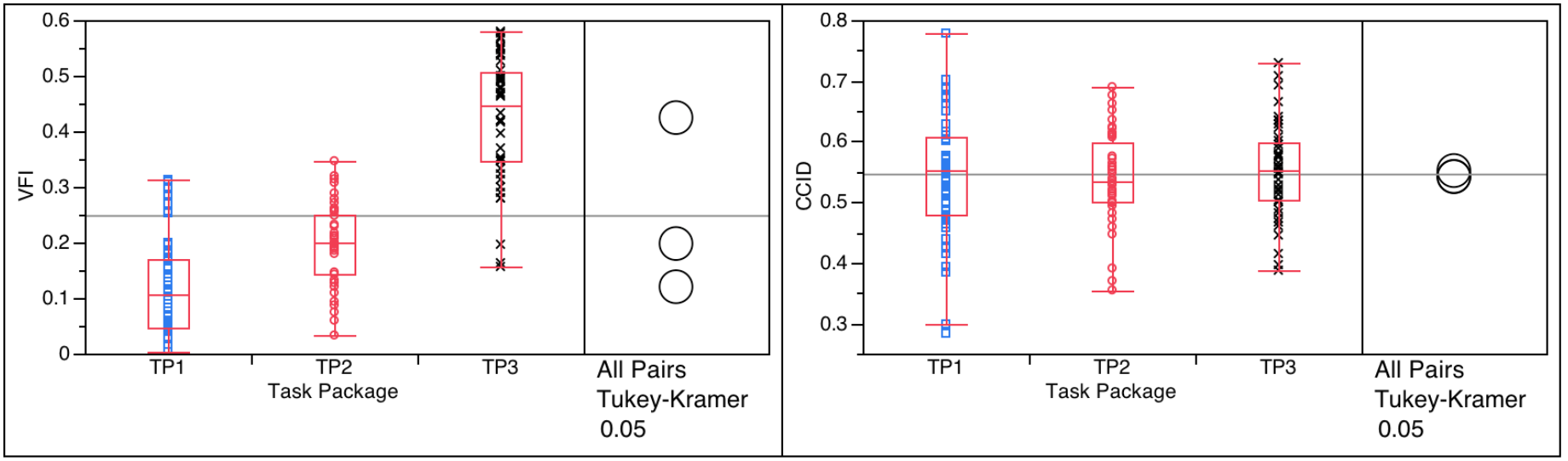
Dependent Variables

- We used two dependent variables, a CCI value for cognitive mode and achievement (r_a)
- We utilized two different methods to create indices of cognitive mode
  - Dunwoody et al. used a linear additive model for CCI_D of five cognitive properties which we emulated
    - judgment control (R_s)
    - kurtosis of error distribution
    - response rate of subjects while making judgments
    - subject self-insight into his judgment policy (subjective)
    - difference in subject confidence in method versus confidence in answer (subjective)
  - We also created a VFI index as an alternate CCI value, in a parallel structure to the VMI index
    - the mean deviation of the matrix product of the cue utilization weights (β_s) with [**R**]

Hypotheses

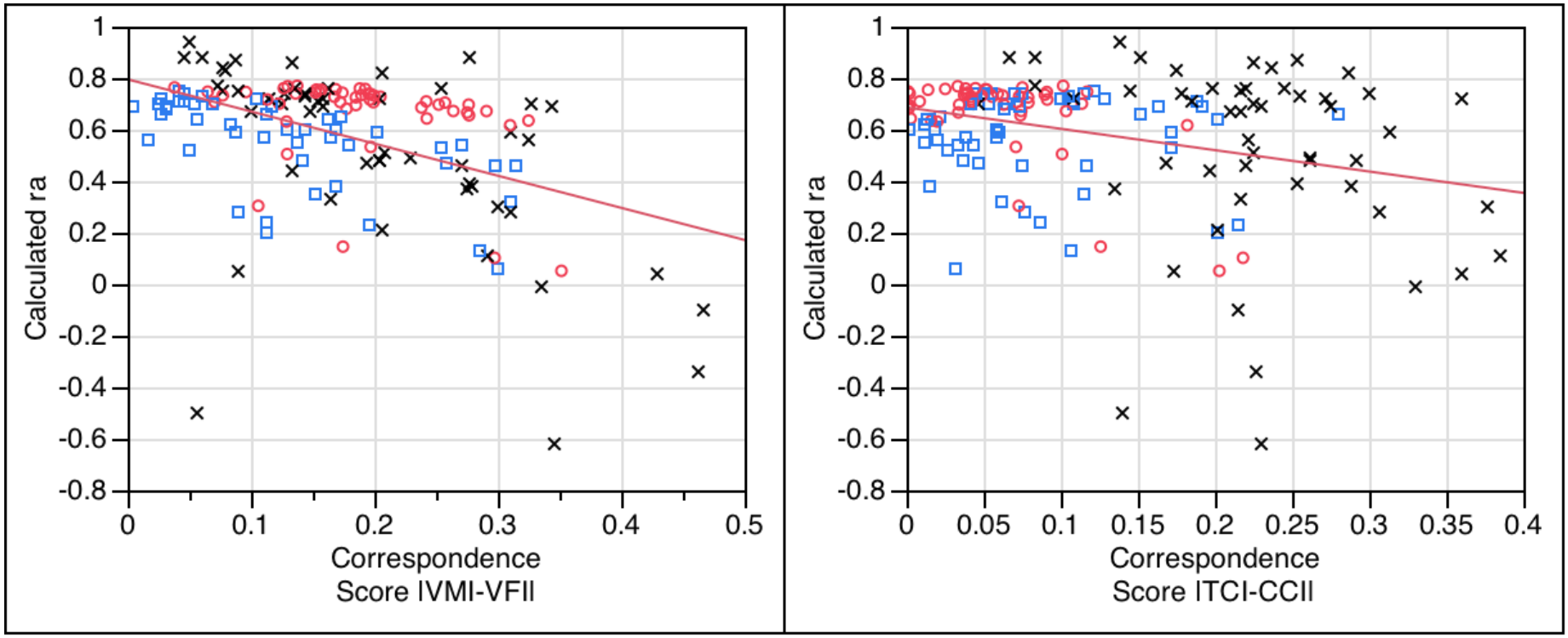
- H_1: The means of the VFI scores will increase as VMI scores increase, reflecting significant shifts in cognitive mode based on task properties. The CCI_D index of cognitive mode will not show similar increases.
- H_2: Achievement will be highest when there is close correspondence between the VMI index and the VFI index (i.e., when task properties match cognitive mode). The TCI_D and CCI_D indices will not show a similar relationship

H_1: VFI Scores by Task Package



- Between subjects analysis
- VFI scores show significant differences and increase as VMI increases, as predicted
- CCI_D scores do not vary as TCI_D increases; no significant difference between them
- This result provides support for CCT because it indicate that as task property index (VMI) increases, cognitive mode index (VFI) also increases

H_2: Achievement Scores (r_a) and Correspondence Score



- VMI-VFI construct shows predicted relationship ($r^2 = 0.22$ (significant at 0.05 level), slope significantly different than zero ($t = -6.64$, $p < 0.001$)) when all tasks considered collectively
- TCI_D-CCI_D construct shows predicted relationship at a lower coefficient of determination ($r^2 = 0.10$ (significant at 0.05 level), slope significantly different than zero ($t = -4.06$, $p < 0.001$)), when all task considered collectively
- VMI-VFI predicted relationship demonstrated in all individual tasks as well as in the aggregate
- TCI_D-CCI_D construct fails to demonstrate the predicted relationship in individual tasks except for TP2

Summary of Experimental Results

- H_1: Task properties induce shift in cognitive mode when measured by VMI-VFI methodology
  - Supported by VFI increases corresponding to VMI increases
  - Not supported by CCI_D metric, which remained unchanged while TCI_D increased
- H_2: Achievement will improve when there is close correspondence between task properties and cognitive mode
  - Supported by VMI-VFI methodology in each task package and in total
  - Only supported by TCI_D-CCI_D methodology in TP2
    - Weakly supported in aggregate for all task packages, but due to bimodal clustering of task package data

Discussion

- Two major facets of CCT were empirically demonstrated in this experiment in contrast to the Dunwoody et al. methodology
  - Support was shown for the relationship between task properties and cognitive mode
  - Support was shown for the relationship between achievement and correspondence between task properties and cognitive mode
- Negatively correlated cues present difficulties for some subjects
  - The impact of negatively correlated cues (and the corresponding cue utilization weight) was seen in the VMI-VFI methodology
  - This effect is masked in the Hammond et al. and Dunwoody et al. methodologies that depend upon average r_{ij} and the standard deviation of β_e.
- Further research into the full impact of negatively correlated cues is warranted in light of these experimental results

Implications of These Results

- These experimental results support the CCT precepts that judgment task properties influence cognitive mode and that close correspondence between the two can result improve judgment performance
  - Using experienced teachers
  - Doing representative tasks
- The result that VMI scores can be predictive of VFI scores can serve to identify optimal cognitive modes for given tasks
  - Which can then serve to identify β_s weighting to maximize achievement potential
  - Knowing the βs weights can be exploited though cue salience in presentation or by choosing the best heuristic for cue utilization weights
- Identifying negatively correlated cues can have significant impact
  - Transforming a negatively correlated cue into a positively correlated one could serve as a mechanism to increase achievement by itself; this hypothesis should be further examined
- The knowledge of task properties from actual case data could serve to accurately predict and therefore improve judgment performance
- If actual case data is lacking then task properties (lens model parameters) can be estimated from other, similar tasks
  - Similar in the sense of similar cue ecological weights and cue inter-correlations
  - Which can then be refined through an iterative process

Next Steps

- Research into CCT has been largely dormant in recent years due to the previous mixed empirical results; this research should revitalize interest in CCT
- The cases presented do not represent the full spectrum of possible cases; further research is warranted into more diverse cases
- Further research is required in the central region of quasi-rationality
  - The poles are easier to define
  - What is the impact of different combinations of intuitive task properties and analytical task properties on achievement?
- Other premises of CCT can now be examined using our methodology
  - Such as dynamic cognition and oscillation between pattern recognition and functional relationships
  - Our metric is a more sensitive tool than the Dunwoody et al. metric