CENTER FOR RESEARCH
ON JUDGMENT AND POLICY

DTIC
S ELECTE D
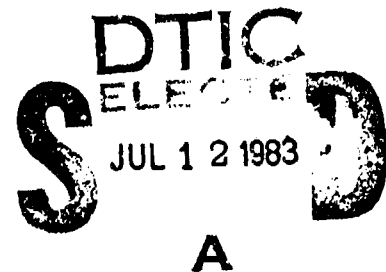JUL 1 2 1983
A

83 07 12 038

DIRECT COMPARISON OF INTUITIVE,

QUASI-RATIONAL AND ANALYTICAL COGNITION

Kenneth Hammond, Robert M. Hamm,

Janet Grassia and Tamra Pearson

Center for Research on Judgment and Policy

University of Colorado at Boulder

Report No. 248

June 1983

DTIC
ELECTE
JUL 1 2 1983

A

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER CRJP 248	2. GOVT ACCESSION NO. A130273	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* Direct Comparison of Intuitive, Quasi-rational and Analytical Cognition		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Kenneth R. Hammond, Robert M. Hamm, Janet Grassia and Tamra Pearson		8. CONTRACT OR GRANT NUMBER(s) N00014-77-C-0336
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Research on Judgment and Policy Institute of Behavioral Science University of Colorado, Boulder, CO 80309		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research 800 North Quincy Street Arlington, VA 22217		12. REPORT DATE June 1983
		13. NUMBER OF PAGES 71
14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)*		15. SECURITY CLASS. *(of this report)* Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; Distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

achievement, cognitive control, confidence, consistency, environmental models, experts, intuitive/analytical, Lens Model, cognitive continuum theory, social judgment theory, task characteristics, transportation, safety

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

The relative efficacy of intuitive and analytical cognition in analytically competent persons was directly compared. More subjects performed best in the intuitive mode when inconsistency was removed from their judgments, an indication that the subjects possessed implicit knowledge that they did not utilize in the analytical mode. More subjects made larger errors in the analytical mode than in the intuitive mode. Subjects' confidence was generally inappropriately placed.

DD FORM 1473   EDITION OF 1 NOV 65 IS OBSOLETE
1 JAN 73
S/N 0102-LF-014-6601

Acknowledgments

DIRECT COMPARISON OF INTUITIVE, QUASI-RATIONAL

AND ANALYTICAL COGNITION

Comparison of the efficacy of intuitive and analytical cognition has been a major topic in judgment and decision research since its inception (for a review of approaches to this topic, see Hammond, McClelland & Mumpower, 1980). Although intuitive cognition is largely believed to be inaccurate, systematically biased, and to produce judgments in which persons are overconfident, (for reviews see Slovic, Fischhoff & Lichtenstein, 1977; Einhorn & Hogarth, 1981), Hastie in his review (1983) concludes that: "none of the applications [of probability theory] to social judgment provide compelling arguments that human reasoning deviates from normative model prescriptions" (p. 511). Our premise is that both conclusions are based on research that is restricted in the following ways:

1. Comparisons are indirect.

The most widely used research method compares the intuitively derived judgments of subjects with analytically derived answers generated by formal models such as Bayes' Theorem and multiple regression equations (see, for example, Kahneman, Slovic, & Tversky, 1982; Hammond, Stewart, Brehmer, & Steinmann, 1975). Informative as these comparisons may be, they do not directly compare the achievement of intuitive and analytical cognition undertaken by the same person. Direct comparisons are needed in

order to know which mode of human cognition is superior. (For a general exposition regarding these comparisons see Tweney, Doherty & Mynert, 1981.)

2. Research subjects lack analytical competence.

Research subjects--typically college sophomores--generally lack knowledge of any analytical principles that would enable them to engage in analytical cognition. Moreover, they are seldom given the time or opportunity to do so. Since the subject's analytical competence is not examined in parallel with his/her intuitive competence, the substitution of formal models to represent analytical cognition constitutes a special comparison. That is, intuitive cognition is set in competition with a form of analytical cognition refined by centuries of study by thousands of individuals over their lifetimes. Although such comparisons are undeniably important, they do not exhaust our range of interest in cognition; they do not inform us about what happens when the same person engages in different forms of cognitive activity. They cannot, for example, inform us about the possible disadvantages as well as advantages of engaging in analytical cognition.

3. Intuition and analysis remain obscure concepts.

Researchers in cognition almost never explain what they mean by intuition, although they take great pains to differentiate precisely among formal, analytical models of cognition. As a result it is customary to explain that intuition is what analysis is not. Brooks (1978), for example, compares "analytic and non-analytic concept formation." Kahneman and Tversky (1982) indicate that "a judgment is called intuitive if it is

reached by an informal and unstructured mode of reasoning, without the use of analytical methods or deliberate calculations" (p. 494). Even philosophers (e.g., Cohen, 1981) who criticize psychological research regarding intuitive judgments of probability fail to say what they mean by this term. As a result, the task properties that induce either form of cognition are not specified, thus making their comparison difficult, if not impossible.

4. **Methodological and substantive cognition are not differentiated.**

Successful cognitive efforts require substantive as well as methodological competence; neither alone can insure empirical success (cf., Hammond, 1966, pp. 68-75; Wason & Johnson-Laird, 1972; Adelman, 1981). Topical knowledge can be as important as the knowledge of methodological principles. These two aspects of cognition are generally studied separately, however. Judgment and decision researchers contrast their subjects' intuitive methodological cognition (efforts to combine information) with analytically derived normative methodological rules for combining information (e.g., Bayes' Theorem), irrespective of subjects' topical knowledge (see Einhorn & Hogarth, 1981; Slovic et al., 1977). Researchers who study problem-solving, on the other hand, generally contrast their subjects' manipulation of the substantive materials of the task with scientific truths or experts' behavior (see, for example, Larkin, McDermott, Simon & Simon, 1980; also Duda & Shortliffe, 1983) without regard for subjects' methodological competence. If the efficacy of intuitive and analytical cognition are to be directly compared, however, these two aspects of cognition need to be examined together, for they can be compensatory in many task circumstances. (An explicit effort to study both can be seen in Fox, 1980.)

5. <u>Evaluations of intuitive methodological cognition are norm-contingent.</u>

Considerable uncertainty exists about which normative rules should be used to evaluate intuitive methodological cognition. The sharply divergent views of philosophers on this topic (see, for example, Kyburg & Smokler, 1980, Kyburg, 1974, and Levi, 1980) are reflected in the responses to Cohen's (1981) criticism of judgment and decision research that uses the standard probability calculus to evaluate intuitive methodological cognition. The number and variety of normative inductive rules cited in these responses as the proper rule to follow are clear evidence that both frequentists and subjectivists have failed to solve the fundamental problem of induction. Until they succeed, the research psychologist who attempts an empirical evaluation of the extent to which intuitive cognition conforms to normative rules will inevitably produce results contingent upon the normative rules s/he prefers--as those who prefer different rules will be quick to point out. Einhorn and Hogarth (1981) cite other difficulties with normative rules in their review when they note that choice of an optimal model is "conditional on certain environmental assumptions and a specific time horizon" (p. 55) and conclude their discussion by observing that "To consider human judgment as suboptimal without discussion of the limitations of optimal models is naive" (p. 56).

Evaluations of substantive analytical cognition are less vulnerable to norm-contingent restrictions, however, because a subject's production of a substantive law or rule can be evaluated with respect to well agreed upon empirical events (cf. Scribner, 1977, who distinguishes between "theoretical" and "empirical" standards for evaluating logical processes). Therefore, empirical achievement is a criterion for cognitive activity

that deserves as much or more interest as conformity with one, among many, normative rules. (See Hammond et al., 1980, on "logical vs. empirical optimality," p. 215ff; see also Einhorn & Hogarth, 1982, on "truth vs. accuracy".)

6. Evaluations of intuitive methodological cognition are task contingent.

Current conclusions about intuitive competence in human beings are largely restricted to the performance of subjects working with methodological problems that have been constructed to be readily treated by the methods of the subjective interpretation of the standard probability calculus. But, as noted above, cognitive problems involve substantive competence as well as methodological competence. Therefore, comparison of the efficacy of intuition and analysis requires examination in tasks that offer challenges to empirical accuracy as well as intuitive methodological competence.

Examination of performance in substantive tasks representative of a person's intellectual environment should be at least as instructive as an examination of performance in task conditions specifically constructed to conform to normative methodological rules. Johnson-Laird and Wason, for example, note in their 1977 "Postscript": "by attempting to relate the...task more closely to the subjects' experience, performance was dramatically improved" (1977, p. 151). Einhorn and Hogarth (1981) conclude: "It is essential to emphasize that the cognitive approach has been concerned primarily with how tasks are represented. The issue of why tasks are represented in particular ways has not yet been addressed" (p. 57). Indeed, the general absence of the description and classification of cognitive tasks has long been emphasized as a serious shortcoming in judgment and decision research (see Hammond, 1954; Edwards, 1971; Slovic

& Lichtenstein, 1971;   Einhorn & Hogarth, 1981) and continues to be
(Hastie, 1983), but without tangible result.

In sum, in addition to making indirect comparisons it is necessary to
directly compare an analytically competent person's use of intuitive and
analytical cognition, to permit the subject to employ either or both in
relation to substantive as well as methodological cognition, and to evaluate
his or her achievement empirically. In the present study, therefore, we do
the following:

1. directly compare intuitive and analytical cognition in the same
    person;

2. use as subjects analytically competent persons so that both intuitive
    and analytical cognition can be brought to bear on the same task;

3. specify the properties of intuitive and analytical cognition and the
    properties of tasks that evoke each;

4. include substantive as well as methodological aspects of cognition;

5. contrast the efficacy of intuition and analysis in terms of empirical
    achievement; and

6. use problems representative of the subject's intellectual environment.

Theoretical Background

A Cognitive Continuum vs. a Cognitive Dichotomy

Our basic premise is that cognitive activity is not a dichotomy of intuition and analysis but rather a continuum marked by intuition at one pole and analysis at the other. Unlike the traditional dichotomous premise, a cognitive continuum permits a quasi-rational compromise between intuitive and analytical cognition; thus a form of cognition more common than either pure intuition or analysis may be described. (For previous use of the concepts of cognitive continuum and cognitive compromise see Brunswik, 1952, 1956; Hammond, 1955; Hammond & Brehmer, 1973; Hammond et al., 1975; Hammond et al., 1980; Hammond, Note 1; see also Anderson, Deane, Hammond, McClelland & Shanteau, 1981.)

Modes of Cognition

Intuition and analysis can be distinguished by the relative degree of: (a) cognitive control (in intuition, low; in analysis, high); (b) rate of data processing (in intuition, rapid, i.e., as brief as microseconds; in analysis, slow, i.e., as long as hours); (c) conscious awareness of process (in intuition, low; in analysis, high); (d) type of organizing principle (in intuition, a weighted average; in analysis, other, task-specific principles); (e) type of error (in intuition, normally distributed; in analysis, few, but large errors); (f) type of confidence (in intuition, confidence in answer but not method; in analysis, confidence in method, not answer). (For further distinctions, see Hammond, Note 1.) The compromise form of cognition, quasi-rationality or "common sense," includes properties from both types of cognition and, therefore, is described in terms of the number and nature of the cognitive properties it includes from both.

Task Properties That Induce Different Modes of Cognition

If, as in the present study, subjects are not provided with feedback, the task properties that differentially induce intuition and analysis include: (a) number of cues available (in intuition, large [> 5]; in analysis, small); (b) the order in which cues are displayed (in intuition, simultaneous; in analysis, sequential); (c) the type of cue measurement required (in intuition, perceptual; in analysis, objective, as with instruments); (d) cue distribution characteristics (in intuition, continuous, highly variable, normally distributed; in analysis dichotomous, valued in terms of specific numbers, distributions unknown); and (e) redundancy among cues (in intuition, high; in analysis, low). (See Hammond, Note 1, for further elaboration.) Quasi-rationality is induced to the extent that tasks contain properties from both types of polar task conditions. We do not claim that all of these properties must be present in order to locate a task at either pole of the continuum, nor do we know their relative importance or their interactive effects that produce quasi rationality. We assert only that specification of these task properties is a useful guide for direct comparison of the relative efficacy of different modes of cognition.

Objectives

Four aspects of the cognitive continuum theory are examined in the context of the six conditions indicated above. Because of the long-standing interest in the relative efficacy of modes of cognition we first examine differences in the empirical accuracy of highway engineers' judgments of safety in the intuitive, quasi-rational and analytical modes. Highway engineers were chosen as subjects because of their frequent professional use of all three modes of cognition. Judgments of highway safety were chosen

because, as we show below, the properties of this task should induce quasi
rationality, thus illustrating the importance of the middle range of the
cognitive continuum. (An intuition-inducing task--judging highway
aesthetics--and an analysis-inducing task--judging highway capacity--were also
employed, but space prevents their discussion here.)

Second, because of the important theoretical role of cognitive
control--predicted to be low in intuition, middle-level in quasi rationality,
and high in analysis--we examine differences in cognitive control in each of
the three modes. More specifically, the relative contributions of knowledge
and cognitive control to accuracy are compared. These two aspects of
cognition, identified by Hammond and Summers (1972), have been found to be
empirically significant in numerous studies of multiple-cue probability
learning (see, e.g., Brehmer, 1979), of interpersonal conflict and
interpersonal learning (see Brehmer, 1972, 1976; Holzworth, in press;
Brehmer & Hammond, 1977 for a review), of the differential effects of various
psychoactive drugs on interpersonal learning and interpersonal conflict among
psychotic patients (Gillis, 1975, 1978), and in clinical judgment (e.g.,
Fisch, Joyce, Hammond & O'Reilly, 1982; Kirwan, Chaput de Saintonge, Joyce, &
Currey, in press). In this study knowledge and cognitive control are examined
to determine whether they provide different contributions to achievement under
the three modes of cognition within the same person.

One of the advantages of the direct comparison of human intuitive and
analytical cognition is that the errors made in the analytical mode of
cognition can be observed and compared with the errors of intuition. This
comparison cannot be made when intuitive cognition is compared only with a
formal model. For in that case the answer provided by the formal model
exhausts the concept of truth; by definition, errors do not exist. Therefore

our third objective is to compare the errors produced by analytical as well as intuitive and quasi-rational cognition.

Brunswik has already argued and demonstrated (1956, pp. 89-93) that intuitive cognition produces errors normally distributed around the correct answer because "intuitive perception must integrate many avenues of approach, or cues...none of which is foolproof or fully ecologically valid" (p. 92). Because analytical cognition proceeds in the opposite fashion, it produces many exactly correct answers, but its errors are likely to be extreme. We therefore ask whether intuitive cognition produces a normal distribution of errors centered on the correct answer; whether analytical cognition produces precisely correct answers together with highly incorrect answers; and whether quasi-rationality produces a distribution of errors lying between those generated by the polar modes of cognition.

Fourth, the relation between confidence and performance is examined to discover whether confidence matches performance (Oskamp, 1965; Lichtenstein, Fischhoff & Phillips, 1982): is an engineer most confident in that cognitive mode in which he performed best and least confident in the mode under which he performed most poorly? The theory outlined above suggests an additional hypothesis. Since the method of intuitive cognition produces rapid, nonretraceable answers and since these answers are based on multiple "avenues of approach or cues," subjects should be less confident in the intuitive method than in intuitive answers. The opposite should be true in the analytical mode, in which the subject's attention is focused on the organization of cues.

Method

Subjects

Twenty-one male highway engineers, 30-70 years of age, served as research subjects. Since engineers are professionally trained to cope with problems that have substantive analytical components, the intuitive and analytical cognitive efforts of the same subject may be directly compared. The engineers' task was to evaluate the safety of highways, a complex problem representative of those encountered in their work.

Independent Variables

In the intuition-inducing condition film strips of one- to three-mile segments of forty two-lane rural Colorado highways were presented. Engineers judged the safety of each segment solely on the basis of the visual material in the film strips (see Figure 1). This form of presentation meets the conditions for inducing intuitive cognition by requiring the engineers to observe a large number of cues contemporaneously displayed, and to measure them by unaided visual perception. The values of the cues are generally continuous and normally distributed; the cues are frequently redundant. The engineers were neither asked nor given the opportunity to organize the task materials explicitly.

Figure 1. View of Highway #15 from Film Strip

In the condition designed to induce quasi rationality,   the   same   forty
highway   segments   were presented as bar-graph profiles which displayed values
for ten dimensions (see Figure 2).   The bar-graph presentation meets the
specifications for inducing quasi-rational cognition by combining intuition-
and analysis-inducing properties.   On the one hand, the task   remains
intuition-inducing because the number of cues is still large; they are
redundant and contemporaneously displayed; and they have continuous,   largely

HIGHWAY JUDGMENT PROJECT 1982

BAR GRAPH PRESENTATION.   SAFETY JUDGMENT TASK.

| | 8   10 | 13 |
| Lane Width | XXXXXXXXXXXXXXX | |

| | 0   4 | 10 |
| Shoulder Width | XXXXXXXXXXXXXXX | |

| | 0   41 | 80 |
| Percent No Passing Zone | XXXXXXXXXXXXXXXXXXXXX | |

| | 0   1 | 7 |
| Curves per Mile | XXXXXX | |

| | 0   47 | 75 |
| Grade | XXXXXXXXXXXXXXXXXXXXXXXXXX | |

| | 0 470 | 10000 |
| Traffic Volume | XXX | |

| | 0   21 | 35 |
| Traffic Mix | XXXXXXXXXXXXXXXXXXXXXXXX | |

| | 0   .7 | 4 |
| Intersections per Mile | XXXXXXX | |

| | 30   42 | 60 |
| Average Speed Limit | XXXXXXXXXXXXXXX | |

| | 0   4.4 | 18 |
| Obstacles per Mile | XXXXXXXXX | |

Figure 2. Bar-graph Profile of Highway #27

normal distributions.   The engineers were not given the opportunity to organize the task materials explicitly and were not asked to explicate or justify their judgments. On the other hand, the number of observable cues is reduced to ten.   The cue values are expressed as numbers, and thus the engineers are not required to measure them perceptually.   The bar-graph profiles also permit easy cross-comparison of the values of various cues.

In the analysis-inducing condition each engineer was requested to construct and justify a mathematical equation for predicting safety (see Appendix A for examples). The subject was instructed to think; pencil, paper and hand calculators were made available;   the engineer was told that a certain amount of time would be allowed for the completion of the formula, but that the time period would be extended until he completed the task. Each engineer was assigned to one of three subgroups within the analytical condition. In the "minimal guidance" subgroup ($n = 12$) engineers were allowed to work through the task with a minimum of instructions and requests from the researchers.   Six engineers in this group were encouraged to complete their formulas in twenty to thirty minutes; the other six were give a target of forty-five minutes.   In the "think aloud" subgroup ($n = 6$) engineers were requested to think aloud as they constructed their equations   so that their analytical efforts could be monitored during the task. In the "maximal guidance" subgroup ($n = 3$) engineers constructed their equations according to a detailed written procedure (see Appendix B) designed to increase the likelihood of a systematic approach to the problem.   (Samples of engineer's remarks regarding their cognitive activity in each mode are presented in Appendix C.)

These three sub-conditions were used because they offered a variety of information. The minimal guidance condition revealed how engineers would go about constructing a formula on their own. The "think aloud" condition provided an understanding of the engineers' process of constructing an equation which would not otherwise have been gained (see Appendix D). The maximal guidance condition showed whether a structural approach was feasible.

It was not our purpose to determine the effects of the different sub-conditions. As would be expected, however, engineers in maximal guidance and engineers who were given a 45-minute target in minimal guidance took more time than the others. There were no significant differences in achievement among the subgroups.

Dependent Variables

The effect of the intuitive, quasi-rational, and analytical conditions were examined with respect to (a) each engineer's degree of achievement in predicting safety accurately, (b) the differential contributions of knowledge and cognitive control to this achievement, (c) the relative frequency of different types of errors made in each of the three modes of cognition, and (d) the relation of confidence to performance in each mode.

Achievement was measured by the correlation between an engineer's estimate of safety and the accident rate of each highway segment (the criterion).

Knowledge and cognitive control were derived from the parameters of the Lens Model Equation (Hammond et al., 1975):

$$r_a = G\, R_e\, R_s + C\sqrt{1 - R_e^2}\ \sqrt{1 - R_s^2}$$

where

r_a = achievement, the correlation between the engineer's judgments and criterion values

G = the correlation between judgments and criterion values corrected for attenuation due to less than perfect linear predictability in each

R_e = environmental predictability (linear form)

R_s = subject's predictability (linear form)

C = correlation between residuals from linear predictions of criterion and residuals from linear predictions of subject's judgments.

In the absence of significant correlations between residuals (trivial values of C in the above equation) then $r_a = G\, R_e\, R_s$. Under these conditions G represents the engineer's knowledge because it indicates what the subject's achievement would have been if he had executed his judgment policy with perfect cognitive control (i.e., $R_s = 1.00$) and if the environmental task criterion were perfectly predictable from the cues (i.e., $R_e = 1.00$).

Cognitive control is appropriately measured by R_s in the equation since there was little evidence in this study of lack of fit of the linear model. (See Hammond et al., 1975, for a detailed discussion of the distinction between consistency and cognitive control.)

Differential errors produced by each engineer are evaluated by (a) examination of the distribution of deviations of judgments around the correct answer, and (b) examination of mistakes made in the process of formulating the equation and evaluation of their consequences for accurate prediction.

Confidence in answers in the intuitive and quasi-rational modes is measured by the mean of the engineer's confidence, on a 1-10 scale, in each of his judgments. Confidence in answers in the analytical mode and confidence in method for all three modes was measured by questions at the conclusion of the session.

Procedure

Statistical Properties of the Task

The same set of forty highway segments was used in all conditions. The criterion for the accuracy of judgments is the accident rate, averaged over seven years, for each highway segment. Accident rate is defined as the total number of accidents (involving fatalities, injuries, or property damage only) divided by the number of vehicle miles traveled. Due to its extremely high accident rate, one highway was dropped from the analysis.

Highways were measured on ten dimensions, chosen for inclusion in the study on the basis of discussions with highway safety experts who indicated the information they considered essential for evaluating the safety of a road (see Table 1 for list). Eight of these measures were available from highway department records; two measures (number of curves per mile and number of obstacles per mile) had to be counted by the experimenters from visual inspection of film strips of each highway segment. The beta weights for each dimension or cue in predicting accident rate are also presented in Table 1.

Visual examination of the scatter-plots of the relations between each cue and

the criterion indicated little if any nonlinear co-variation. This finding

was supported by the results from calculation of the contribution of squared

terms and interactions to accident rates.

Table 1

Highway Characteristics (cues) Related to Accident Rate

Cues	Beta weights
Lane Width	.023
Shoulder Width	-.042
Percent No Passing Zone	-.143
Curves per Mile	.152
Grade	.055
Traffic Volume	-.198
Traffic Mix (% of Trucks)	-.017
Intersections per Mile	.247
Average Speed Limit	-.316
Obstacles per Mile	.478

An optimally weighted linear multiple regression model of the task indicates that R_e = .863, corrected for estimated shrinkage = .809. Application of equal weights and linear functions, each cue given the sign that appeared in the best fit equation, yields an R_e of .769 (corrected for shrinkage = .667). The intercorrelations among the ten cues and the criterion are presented in Table 2.

Rating Scales

The judgment scale for each task condition was appropriate to the cognitive activity induced. An abstract rating scale from 1 (safe) to 10 (unsafe) was used for the film strips to induce intuitive cognition. In the bar-graph presentation and in the task requiring the construction of a formula, a scale from 0 to 32 accidents per million vehicle miles traveled was employed because this specificity is compatible with calculation and thus with analytical cognition. Transformations to a common scale were made for purposes of data analysis and are described below.

Trials

All engineers were presented with the tasks in the same order: first, the film strips; second, the bar graphs; third, the materials for formula construction. It was not appropriate to counterbalance the order of presentation because analytical work requiring use of certain cues in an explicit fashion would have strongly influenced subsequent intuitive judgments, whereas the reverse is not true (see Jones and Harris, 1982).

In the intuition-inducing mode, ten of the forty highways were shown twice; in the quasi rationality-inducing mode sixteen highways were shown twice. These repetitions permitted calculation of repeated trials reliability for each engineer.

Table 2

Intercorrelations among Cues

	LWIDTH	SWIDTH	PCTNPZ	CURVEPM	GRADE	TRAFVOL	TRAFMIX	INTPM	AVESL	OBSPM
SWIDTH	.414									
PCTNPZ	-.046	-.251								
CURVEPM	.203	-.357	.554							
GRADE	-.294	-.254	.152	.360						
TRAFVOL	.492	.214	.135	.169	-.098					
TRAFMIX	-.229	.037	-.112	-.280	-.009	-.330				
INTPM	-.070	-.076	-.010	-.284	-.159	-.083	.032			
AVESL	.101	.457	-.481	-.591	-.265	-.106	.322	-.165		
OBSPM	.068	-.408	.647	.747	.291	.148	-.315	.079	-.674	
TRATEAV	-.083	-.466	.393	.558	.295	-.105	-.230	.304	-.717	.743

Time

Response times were recorded in all three presentation modes, but are artifacts of the task conditions rather than descriptions of the subjects' behavior; that is, presentation of a film strip required more time than presentation of a bar graph. The mean response time in the intuitive mode was 78 seconds and in the quasi-rational mode, 22 seconds.

In the analytical mode the response time varied within the three subgroups. The eighteen engineers in the think aloud and minimal guidance subgroups were encouraged to complete their formulas in twenty to forty-five minutes; but seven of them took at least an hour at this task (maximum = 115 minutes). No time constraints were imposed on the engineers in the maximal guidance condition; response times ranged from 138 to 250 minutes with a mean of 180 minutes.

RESULTS

Relative Efficacy of Three Modes of Cognition

Achievement

Individual differences in achievement. Achievement (r_a) is measured by the correlation between an engineer's judgments about the safety of each of the 39 highways and the accident rates of those highways. Data from repeated judgments in the intuitive and quasi-rational modes were not included. For the analytical condition, the engineer's formula was applied to each of the 39 highways, and the answers thus produced were correlated with the accident rates.

Wide individual differences occurred in all modes of cognition (Table 3).
In the intuitive mode, the engineer's achievement correlations ranged from
.071 to .636, with a median of .576. In the quasi-rational mode, r_a ranged
between .000 and .738, with a median of .516. In the analytical mode, ra's
ranged from -.226 to .731, with a median of .467. Mean interjudge agreement
in the intuitive mode is .58; in the quasi-rational mode, .47; and in the
analytical mode, .17. The correlation between engineers' achievement in the
intuitive and quasi-rational modes is .344; between achievement in the
intuitive and analytical modes, .457; and between achievement in the
quasi-rational and analytical modes, .032 (Table 4).

Since the engineers' general policies for judging highway safety in the
intuitive and quasi-rational modes were all modeled by best-fit multiple
regression analysis, individual differences in achievement in these modes
cannot be attributed to the structure of this single model. In the analytical
mode, however, judgment policies were not modeled but instead directly
expressed by the engineers as formulas. Individual differences in achievement
in the analytical mode could have been caused, in part, by the wide variety in
the structure of those formulas (see Appendix A). It turned out, however,
that the various structural features of the formulas are not related to
achievement.

In the intuitive mode, the more experience the engineer had (measured by
age, years of work, or years of education), the lower his cognitive control,
R_s; otherwise there were no significant relations between experience and the
various measures of achievement, knowledge and consistency. The existence of
wide individual differences in achievement among the engineers in all three
cognitive modes thus supports the decision to study separately the performance

Table 3

Lens Model Equation Components and Repeated Trials Reliabilities

for Engineers in Each Condition

ENGINEER NUMBER	Achievement (r_a)			G			Cognitive Control (R_S)			Repeated Trials Reliability		
	I	Q	A	I	Q	A	I	Q	A	I	Q	A
1	.378	.377	.265	.517	.545	.307	.818	.690	1.000	.870	.229	1.000
2	.625	.712	.071	.770	.845	.082	.790	.878	1.000	.612	.652	1.000
3	.071	.436	-.226	.141	.534	-.262	.730	.726	1.000	.848	.484	1.000
4	.606	.624	.161	.825	.786	.187	.891	.929	1.000	.860	.925	1.000
5	.636	.609	.641	.946	.745	.743	.795	.914	1.000	.759	.917	1.000
6	.500	.499	.518	.646	.644	.600	.694	.717	1.000	.705	.694	1.000
7	.521	.666	.581	.731	.778	.673	.693	.919	1.000	.896	.957	1.000
8	.576	.446	.731	.836	.639	.847	731	.856	1.000	.848	.836	1.000
9	.378	.120	.436	.523	.291	.505	.756	.737	1.000	.716	.200	1.000
10	.632	.227	-.043	.873	.324	-.056	.785	.665	1.000	.848	.394	1.000
11	.577	.738	.296	.910	.863	.343	.644	.945	1.000	.834	.894	1.000
12	.572	.416	.524	.783	.534	.6C7	.792	.836	1.000	.938	.789	1.000
13	.472	.000	.390	.725	-.007	.452	.624	.571	1.00C	.618	.387	1.000
15	.473	.518	-.026	.742	.721	-.030	.707	.785	1.000	.809	.561	1.000
17	.612	.691	.436	.838	.864	.505	.777	.886	1.000	.980	.805	1.000

Table 3 (Continued)

Lens Model Equation Components and Repeated Trials Reliabilities

for Engineers in Each Condition

ENGINEER NUMBER	Achievement (r_a)			G			Cognitive Control (R_s)			Repeated Trials Reliability		
	I	Q	A	I	Q	A	I	Q	A	I	Q	A
19	.589	.516	.681	.795	.574	.789	.743	.843	1.000	.891	.659	1.000
20	.618	.462	.670	.816	.634	.776	.767	.851	1.000	.754	.402	1.000
21	.583	.538	.691	.899	.806	.801	.756	.895	1.000	.771	.777	1.000
22	.636	.563	.643	.835	.780	.745	.778	.893	1.000	.871	.824	1.000
23	.437	.591	.467	.636	.852	.541	.594	.887	1.000	.958	.838	1.000
24	.572	.341	.636	.724	.468	.737	.794	.871	1.000	.751	.799	1.000
$\bar{X}_r$*	.536	.500	.438	.777	.572	.528	.754	.846	1.000	.845	.735	1.000

* Based on z transformations.

Table 4

Intercorrelations among Lens Model Components

	Achievement (r_a)			G			Cognitive Control (R_s)			Repeated Trials Reliability		
	I	Q	A	I	Q	A	I	Q	A	I	Q	A
r_{IQ}	← .344 →			← .372 →			← .167 →			← .319 →		
r_{QA}		← .032 →			← .052 →			---			---	
r_{IA}	← .457 ——		→	← .385 ——		→	---			---		

of each engineer (cf. Brunswik, 1952; Luce, 1959; Tversky, 1972; Shanteau, 1975; Hammond et al., 1980, pp. 115-127; Hammond & Wascoe, 1980; Hammond, Note 1).

Achievement in three modes of cognition. In order to make direct comparisons of each subject's performance in each mode of cognition, the efficacy of each mode for each engineer was evaluated. The basic analysis of data thus consisted of rank ordering each engineer's performance across modes. Means are provided for general information but are not included in the statistical tests.

A chi-square test of the hypothesis that one mode of cognition was superior to the others resulted in a failure to reject the null hypothesis (see Table 5). There were, however, only four engineers whose achievement was highest in the intuitive mode. Comparison of the six orders of achievement did result in rejection of the hypothesis of a chance distribution (p < .01). The data suggest that all nine engineers who performed best in the analytical mode performed better in the intuitive than the quasi-rational mode.

Components of achievement. As the Lens Model Equation indicates, achievement (r_a) is a function of knowledge (G) and cognitive control (R_s). Since two subjects might have equal achievement for different reasons--that is, their knowledge and their cognitive control over the application of their knowledge might vary in a compensatory manner--we determined the relative contribution of these two components for each engineer in each cognitive mode.

Table 5

Number of Engineers with Highest Achievement (r_a) in Each Condition

	1st	Order	
I	4	I > Q > A	2
		I > A > Q	2
Q	8	Q > I > A	6
		Q > A > I	2
A	9	A > I > Q	9
		A > Q > I	0
	21		21

$$\chi^2 = 2.00, \; p > .20 \qquad \chi^2 = 15.86, \; p < .01$$

For the intuitive and quasi-rational data sets, the values of the parameters of the Lens Model Equation were calculated from the best fit linear models of the environment and of each engineer's judgments. For the analytical data set, the best fit linear model was used for the environmental system, but the engineer's judgment policy was represented by the formula he produced. In some cases the formulas had nonlinear features. An additional

analysis was carried out using a linear best fit model of the answers generated form the engineer's formula; the results differed only slightly from those produced by the first procedure and will not be reported here.

Knowledge. Table 6 shows that the knowledge component, G, for thirteen of the twenty-one engineers is higher in the intuitive mode than in the quasi-rational and analytical modes (X^2 = 8.86; p < .02). Examination of order also shows a distribution significantly different from chance (p < .01). No engineer has poorest knowledge in the intuitive mode. Thus we conclude that if the attenuation of r_a by the inconsistency in the task and in each subject's performance were eliminated, most engineers would have achieved higher accuracy of prediction in the intuitive mode than any other. Most engineers would have judged least accurately in the analytical mode. The number of engineers whose quasi-rational judgments were best would have been about the same as the number whose quasi-rational judgments were poorest.

Cognitive control. The cognitive control component, R_s, is of course significantly higher (p < .001) in the analytical mode than in the other modes because the answers in the analytical mode were mechanically generated from the engineers' formulas. Far more engineers had a larger R_s in the quasi-rational mode than in the intuitive mode (16 of 21; p < .05). The relation between cognitive control and repeated trials reliability is discussed in Appendix E.

Table 6

Number of Engineers with Highest Knowledge (G) in Each Condition

	1st	Order	
I	13	I > Q > A	9
		I > A > Q	4
Q	6	Q > I > A	6
		Q > A > I	0
A	2	A > I > Q	2
		A > Q > I	0
	21		21

$$\chi^2 = 8.86, \ p < .02 \qquad \chi^2 = 18.14, \ p < .01$$

Summary of findings regarding achievement. Fewest engineers were most accurate (r_a) in the intuitive mode, although this finding is not statistically significant. Comparison of knowledge as measured by G, however, indicates that even though the engineers explicated their knowledge and applied it rigorously, i.e., with perfect consistency, in the analytical mode, their achievement would have been higher in the intuitive and quasi-rational

modes if they had been equally consistent there.   This result supports the
conclusion that engineers possess and can apply implicit knowledge about
highway safety that they do not accurately express in rigorous, retraceable,
quantitative form.

Differential Distributions of Errors under Different Cognitive Modes

Investigating the covariation between judgment and criterion does not
exhaust the possible methods for evaluation of performance.   Brunswik
suggested (1948, 1956, pp. 89-91; see also Hammond, Note 1) that the use of
different modes of cognition may lead to similar achievement but to very
different types of errors. Based on the distinction between intuition and
analysis described by Brunswik and Hammond, we test the following hypotheses:

1. Errors made in the intuitive mode are normally distributed and also
    less frequently exactly correct and less frequently widely incorrect
    than errors made in the analytical mode;

2. The quasi-rational mode produces an error distribution that lies
    between the distribution of errors in intuitive and analytical
    cognition.

Rescaling procedure.   The error scores were produced by subtracting the
accident rate for a highway from the engineer's judgment. The response scale
used in the intuitive mode ran from 1 (= unsafe) to 10 (= safe), whereas the
scales in the quasi-rational and analytical mode ran from 0 to 32 accidents
per million vehicle miles traveled.   To make the error distributions

comparable, judgments made in the intuitive mode were rescaled onto the 0 to 32 scale. Further, some engineers' analytical formulas did not produce numbers between 0 and 32. Their responses were rescaled by mapping the number they intended their formula to produce for the safest road onto zero, and the number they intended for the most dangerous road onto 32.

Hypothesis 1: The error distributions in the analytical condition should deviate from normal more often than they do in the intuitive condition.

Procedure. For each engineer and for each mode, the mean and standard deviation of each engineer's error distribution were used to produce a normal distribution. The null hypothesis is that the observed distribution of an engineer's errors does not differ significantly from the constructed normal distribution. Therefore, the normal distribution was divided into six categories expected to have equal numbers of observations (the boundaries were located at -.97, -.43, 0, +.43 and +.97 standard deviations away from the mean), and the number of the errors that fell into each category was observed.

Results. In the intuitive, quasi-rational and analytical modes there were few engineers (two, five, and three out of twenty-one, respectively) whose distributions of deviations differed significantly from normal. Additionally, a comparison of the chi-squares for the deviations of these error distributions from normal revealed that no mode tended to have larger deviations than the other modes. These results offer no support for the hypothesis that the analytical mode would produce more non-normal distributions than the intuitive mode.

Hypothesis 1 also asserts that more of the engineers' analytical judgments than intuitive judgments should have error distributions with positive kurtosis; the distributions of errors from quasi-rational judgments should lie between those of the other modes.

Procedure. Since a positive kurtosis indicates a peaked frequency distribution, i.e., one in which the answers are often nearly correct yet occasionally very far off, we examined the degree of kurtosis in each engineer's error distribution.

Results. The results in Table 7 indicate that the error distribution for the analytical mode was more peaked than for the intuitive mode for seventeen of the 21 engineers ($p < .01$). The quasi-rational error distribution was more peaked than the intuitive error distribution for seventeen engineers ($p < .01$). However, for only ten of the engineers was the analytical error distribution more peaked than the quasi-rational. The results from examination of kurtosis thus support the hypothesis that more engineers would produce answers that were less frequently exactly correct in the intuitive mode than in the analytical mode. Finding that the quasi-rational error distribution was not different from the analytical was unexpected.

A second way of testing the hypothesis concerning the shape of error distributions in each cognitive mode is to look at each engineer's performance as a whole rather than at each answer individually. If the formula is correct, achievement (r_a, the correlation between answers and accident rate) should be relatively high because the formula should generate more answers that are nearly correct than intuitive judgments do. If an engineer makes an error in a formula, all answers should be wrong and r_a should be relatively low.

Table 7

Number of Engineers with Various Pairwise Orders of Kurtosis of Error Distribution

A > I	17	$X^2 = 6.86$
I > A	4	$p < .01$
A > Q	10	$X^2 = 0.00$
Q > A	11	NS
Q > I	17	$X^2 = 6.86$
I > Q	4	$p < .01$

In Table 8 the range of correlations (r_a) between answers and accident rates for the five best performances in the analytical mode (.643 to .731) and for the five best performances in the intuitive mode (.618 to .636) show no overlap. From this standpoint the best analytical performances are clearly superior to the best intuitive performances. However, the median achievement in the analytical mode was .467, far worse than the median achievement in the intuitive mode (.576); and the analytical achievements covered a total range of .957 as compared to the intuitive range of .565. Engineers' performance data as well as analysis of error on individual judgments thus supports the hypothesis that analytical cognition is more often very precise yet more often widely in error.

Table 8

Achievement in Each Mode

	Intuitive	Quasi-rational	Analytical
Span of Best Five Engineers' Achievements	.618 - .636	.624 - .738	.643 - .731
Median Achievement (all 21 engineers)	.576	.516	.467
Range of Achievement (all 21 engineers)	.565	.738	.957

Experimenter's Observations Regarding "Widely Incorrect" Answers.

The above analyses describe the distributions of errors produced by different modes of cognition; but the "catastrophic" errors that Brunswik predicted would occur in the analytical mode are best illustrated by examples of errors made by three engineers in the analytical mode:

1. Engineer #2 made a careless arithmetic error in producing weights in his formula. He first assigned a weight of .10 to each of the ten cues. Next he adjusted the weights of important cues to .12. Finally, intending to assign weights of .08 to cues he felt were slightly less important, he wrote instead .8. Thus he gave the highest weight to the cues to which he wished to give least weight. The effects of his error were "catastrophic": his achievement (r_a) was .071, and his mean error was 44.176 on a scale he intended to go from 0 to 32.

2. Although provided with the information that certain cues could assume a value of zero, Engineer #4 produced a formula in which some of these cues appeared in the denominators of fractions. As a result, his answer was indeterminate on 5 of the 39 highways.

3. Engineer #10 produced a linear weighted average formula, in which he inadvertently used the wrong sign on two cues. His achievement was consequently a very low -.048. Had he not made this error his achievement would have been .293.

Since the processes of intuitive and analytical cognition cannot be directly observed, we cannot explain occasional poor performance in the film-strip and bar-graph tasks by citing the kind of careless, one-time errors in execution of a complex symbolic procedure as we did above. It is unlikely that this kind of error is occurring in the intuitive and quasi-rational modes, however. Even if, in the film-strip and bar-graph tasks, the engineers were actually engaging in a step-by-step cognitive procedure similar to that observed in the analytical mode, they would have to make the same error during each trial in the task. Although a careless error is guaranteed to occur with complete consistency when a faulty formula is applied to thirty-nine highways, it is not guaranteed by the conditions of intuitive and quasi-rational cognition. On the contrary, since no engineer made perfectly consistent judgments in either the intuitive or the quasi-rational modes, none could be expected to make consistent errors. Poor performance in the intuitive and quasi-rational modes must instead be attributed either to an engineer's lack of knowledge or to his inability to bring his knowledge accurately to bear on the task.

Summary regarding differential errors. When the distribution of errors
was examined, the results provided no support for the general hypothesis that
errors would be found to be normally distributed in the intuitive mode more
often than in the analytical mode, with errors in the quasi-rational mode
intermediate. However, the predicted differences in kurtosis were observed:
engineer's performances in the analytical mode were more often more accurate
than their performances in the intuitive mode, but occasional large errors
were produced in the analytical mode. Results in the quasi-rational mode were
little different from the analytical mode in this regard. Inspection of the
engineers' analytical work revealed instances of large errors as anticipated
(misplaced decimals, division by zero, and reversed signs). No similar errors
were found in the other modes of cognition.

One reason that the anticipated differences in the degree of normality of
error distributions were not found may be that the formulas produced by the
engineers were quasi-rational in form; that is, many engineers developed
formulas that were analogous to a robust weighted average (eleven by a strict
criterion, twenty by a loose criterion; see Appendix A). Formulas of this
type (when applied without errors in calculations) are as unlikely to produce
large errors as the weighted averages that the evidence suggests the engineers
were using in the intuitive and quasi-rational modes.

Confidence

Answer Confidence Compared with Method Confidence

Cognitive continuum theory predicts that subjects will be more confident in answers produced by intuitive cognition than in answers produced by analytical cognition. Conversely, the theory predicts more confidence in the method of analysis than in the method of intuition. These predictions may be tested by analyzing (a) the order of engineers' confidence among modes of cognition, for answer and method confidence separately, and (b) the order of engineers' answer confidence and method confidence within each mode. These two approaches are diagrammed in Figures 3 and 4.
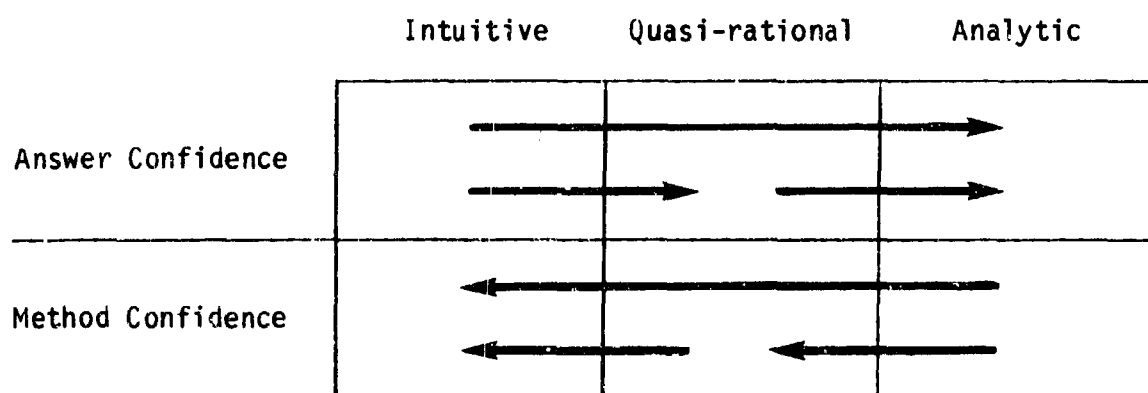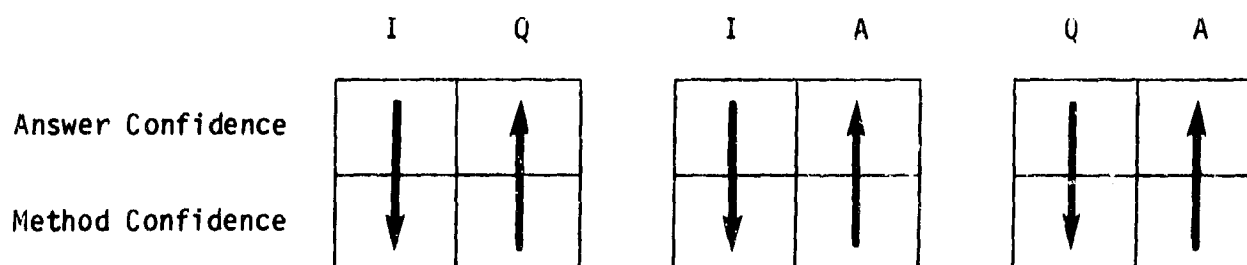

Note: The direction of the arrows indicates decreasing confidence.

Figure 3. Predicted Relative Confidence for Each Pair of Cognitive Modes

Figure 3 illustrates the expected order of answer confidence and of method confidence among modes of cognition. Answer confidence should be relatively high in the intuitive mode and decrease as cognition becomes more analytical. The reverse should be true for method confidence.

Figure 4 represents the expected within-mode order of engineers' answer confidence and method confidence. In each pair of modes engineers should express more method confidence than answer confidence for the mode closer to the analytical pole on the cognitive continuum.

Note: The direction of the arrows indicates decreasing confidence.

Figure 4.   Predicted Relationship Between Answer Confidence and Method
            Confidence for Each Pair of Cognitive Modes

Procedure.

All confidence ratings in this study were indicated on a 1-to-10 scale where 10 = high confidence. In the intuitive and quasi-rational modes, answer confidence is indicated for each engineer by the mean of the confidence ratings he made for each of his judgments. Answer confidence in the

analytical mode and method confidence for all three modes are represented by the means of responses to particular confidence questions (see Appendix F) which the engineer answered after making judgments or constructing a formula.

Results

Orders of Answer and Method Confidence. The rank order of engineers' confidence across modes is shown in Table 9. For answer confidence the modal order (10 engineers) was exactly the predicted order (see Figure 3): more subjects indicate more answer confidence in the intuitive mode than in the quasi-rational mode, and more in the quasi-rational than in the analytical mode. As predicted, twenty engineers had more answer confidence in the intuitive than in the quasi-rational mode ($p < .001$); and sixteen had more in the intuitive than in the analytical mode ($p < .05$); however, only eleven had more answer confidence in the quasi-rational than in the analytical mode (NS).

For method confidence, the modal confidence order was $I > A > Q$, which was not predicted. Thirteen and a half engineers (ties counted as 1/2 in each of the indicated orders) had more method confidence in the analytical mode than in the quasi-rational mode (NS, $p < .20$); but only three had more method confidence in the analytical than the intuitive mode ($p < .01$ in the wrong direction), and only 3.5 had more method confidence in the quasi-rational than in the intuitive mode ($p < .01$ in the wrong direction).

Table 9

Count of Engineers Who Had Each Possible Confidence

Order among the Three Modes

	Answer Confidence	Method Confidence
I > Q > A	10	3
I > A > Q	5	11.5*
Q > I > A	1	2.5*
Q > A > I	0	1
A > I > Q	5	2
A > Q > I	0	0

$$x^2 = 22.14, \; p < .001 \qquad x^2 = 25.78, \; p < .001$$

* When an engineer had equal confidence in two modes, .5 count
was assigned to each of the two orders. For method confidence,
there were two engineers who had ties. Method Confidence data
was missing for one engineer.

Method Confidence Compared with Answer Confidence

Comparison of each engineer's method confidence with his answer
confidence shows that the majority of engineers had greater confidence in
answer than in method for the intuitive and quasi-rational modes and greater
confidence in method than in answer for the analytical mode (see Table 10).
Chi-squared tests showed, however, that this finding is not significant.

Table 10

Relative Answer and Method Confidence for Each Cognitive Mode*

	I	Q	A
Answer Confidence is greater	12	15	7
Equal			2
Method Confidence is greater	8	6	12

*Due to missing data, N = 20 for the intuitive task.

In summary, the prediction that the engineers would have more confidence in their answers in the intuitive mode was born out, but the prediction that the engineers would have more confidence in their method in the analytical mode was contradicted. In the comparison of answer and method confidence within mode, the predicted pattern of greater method confidence than answer confidence in the analytical mode and greater answer confidence in the intuitive mode was observed, but did not reach statistical significance. Most engineers had greater confidence in the intuitive mode for both answers and method. Judging safety was less susceptible to analytical cognition than we had anticipated.

Appropriate Placement of Confidence

Appropriate placement of confidence can be evaluated in terms of whether experts' confidence in their judgments reflects their achievement--that is, are they more confident in the mode in which their achievement is higher?

Table 11 displays for both answer confidence and method confidence the relation between relative confidence and relative achievement for each pair of modes (intuitive versus quasi-rational, intuitive versus analytical, and quasi-rational versus analytical).

In each of the four-celled blocks, the upper left and lower right cells contain the number of engineers whose confidence was appropriate given their achievement; the upper right and lower left cells, engineers whose confidence was misplaced. For example, consider the block relating relative answer confidence to relative achievement for the intuitive and quasi-rational modes. Of the thirteen engineers who had higher intuitive achievement than quasi-rational achievement, twelve of them had appropriate answer confidence; that is, they had higher answer confidence in the intuitive mode than in the quasi-rational mode. However, none of the eight engineers who achieved higher in the quasi-rational mode than the intuitive mode had appropriate confidence.

Table 11

Appropriateness of Confidence Rating to Achievement

a. Answer Confidence

	CONFIDENCE I > Q	CONFIDENCE Q > I
ACHIEVEMENT I > Q	12	1
ACHIEVEMENT Q > I	8	0

$X^2 = 13.67,\ p < .001$

	CONFIDENCE I > A	CONFIDENCE A > I
ACHIEVEMENT I > A	7	3
ACHIEVEMENT A > I	9	2

$X^2 = 3.82,\ p < .10$

	CONFIDENCE Q > A	CONFIDENCE A > Q
ACHIEVEMENT Q > A	2	8
ACHIEVEMENT A > Q	9	2

$X^2 = 0.20,\ \text{NS}$

b. Method Confidence*

	CONFIDENCE** I > Q	CONFIDENCE** Q > I
ACHIEVEMENT I > Q	8	3
ACHIEVEMENT Q > I	8	0

$X^2 = 7.29,\ p < .01$

	CONFIDENCE I > A	CONFIDENCE A > I
ACHIEVEMENT I > A	8	1
ACHIEVEMENT A > I	9	2

$X^2 = 7.27,\ p < .01$

	CONFIDENCE** Q > A	CONFIDENCE** A > Q
ACHIEVEMENT Q > A	8	5
ACHIEVEMENT A > Q	1	5

$X^2 = 1.76,\ \text{NS}$

*Method Confidence information was missing for one engineer; therefore N = 20.
**One engineer expressed equal confidence in both modes; therefore N = 19.
***Expected frequencies were calculated for each cell by A times B times N, where A = the proportion of engineers who had higher achievement in the given mode, B = the overall proportion of engineers who had appropriate or inappropriate confidence, and N = the total number of engineers (excluding ties).

Both answer and method confidence revealed a pattern of overconfidence in the intuitive mode as compared to the quasi-rational mode ($p < .001$ and $p < .01$, respectively). Most engineers also had inappropriately high answer confidence ($p < .10$) and method confidence ($p < .01$) in the intuitive mode when compared to the analytical mode. No pattern of overconfidence was seen between the quasi-rational and analytical modes, although the engineers were very inaccurate. In fact, a null model of random confidence judgments was rejected here because the order of engineers' performance is the opposite of the order of their confidence when the quasi-rational and analytical modes are compa2ed.

Since the engineers received no feedback about their achievement (r_a), the appropriateness of their answer confidence and method confidence must also be evaluated in terms of knowledge (G). The same pattern of overconfidence in the intuitive mode that was observed with achievement is seen with knowledge (Table 12).

Summary Regarding Appropriateness of Confidence

When appropriateness of confidence is evaluated in terms of empirical achievement (r_a) most engineers were found to be (a) more confident about their answers and also about the method used to produce their answers in the intuitive mode than in the other two modes; and (b) overconfident in intuitive cognition and underconfident in analytical and quasi-rational cognition. The general pattern of overconfidence in intuitive cognition and underconfidence in analytical cognition remains when the engineers' knowledge (G) is examined.

Table 12

Appropriateness of Confidence Rating to Knowledge

a. Answer Confidence

KNOWLEDGE \ CONFIDENCE	I > Q	Q > I
I > Q	14	1
Q > I	6	0

$x^2 = 12.86$, p < .001

KNOWLEDGE \ CONFIDENCE	I > A	A > I
I > A	14	5
A > I	2	0

$x^2 = 1.73$, p < .20

KNOWLEDGE \ CONFIDENCE	Q > A	A > Q
Q > A	5	9
A > Q	6	1

$x^2 = 0.26$, NS

b. Method Confidence*

KNOWLEDGE \ CONFIDENCE**	I > Q	Q > I
I > Q	11	3
Q > I	5	0

$x^2 = 6.39$, p < .02

KNOWLEDGE \ CONFIDENCE	I > A	A > I
I > A	15	3
A > I	2	0

$x^2 = 2.96$, p < .10

KNOWLEDGE \ CONFIDENCE**	Q > A	A > Q
Q > A	3	10
A > Q	3	3

$x^2 = 0.41$, NS

*Method Confidence information was missing for one engineer; therefore N = 20.
**One engineer expressed equal confidence in both modes; therefore N = 19.
***Expected frequencies were calculated for each cell by A times B times N, where A = the proportion of engineers who had higher achievement in the given mode, B = the overall proportion of engineers who had appropriate or inappropriate confidence, and N = the total number of engineers (excluding ties).

Summary and Discussion

In an effort to broaden and deepen comparisons of the efficacy of intuitive and analytical cognition six criteria were specified: (a) direct comparisons should be made between intuitive and analytical cognition; (b) the subjects used in the study should be analytically competent by virtue of their professional training; (c) the concepts of intuition and analysis, as well as the task properties that induce each, should be given clear meaning; (d) substantive and methodological cognition should be included in the same study; (e) conclusions regarding the relative efficacy of these modes of cognition should be norm-independent, and thus not subject to disputed standards of inductive inference; and (f) conclusions should not be restricted to problems involving the logic of the standard probability calculus.

These criteria were met by (a) requiring each subject to engage in intuitive and analytical cognition with regard to the same problem, (b) employing professional highway engineers as subjects, (c) specifying the principal characteristics of intuitive and analytical cognition and the properties of the tasks that induce them, (d) distinguishing between methodological and substantive cognition, (e) evaluating performance in terms of empirical achievement, and (f) employing problems that do not require the use of the conventional probability calculus.

The results indicate that within these broadened conditions:

1. Achievement in a cognitive mode is greatly influenced by the degree of consistency present. The absolute consistency of analytical cognition affords it an advantage in empirical achievement over intuitive and quasi-rational cognition, even when the substantive knowledge that is employed analytically is relatively less correct. Conversely, the inconsistency (low cognitive control) of intuitive cognition will lead to an underestimate of its empirical value, unless such inconsistency is identified and removed.

2. The knowledge implicit in intuitive judgments was found to be empirically superior to quantified, formalized knowledge. When inconsistency was removed from the engineers' intuitive and quasi-rational judgments, intuition and quasi rationality were more efficacious than analysis for most of the subjects.

3. Analytical cognition was more often highly accurate, yet more often very inaccurate when compared with intuitive cognition for most of the subjects. Large errors were often made in the analytical mode, but seldom in the intuitive or quasi-rational modes.

4. When performance is evaluated in terms of achievement, engineers were most confident in the intuitive mode, not only in their answers (as anticipated) but in their method as well (not anticipated). However, this greater confidence in the intuitive mode was not always warranted.

These results demonstrate that a cognitive continuum theory has positive utility: it enables the findings of this study concerning achievement, knowledge, errors, and confidence to be largely predicted from the nature of the task properties employed. This conclusion has implications for cognitive research as a whole. Research ranging from the intuition-inducing tasks of judgment and decision making to the analysis-inducing tasks of problem solving may also be understood to be located along a cognitive continuum. Specification of task properties and description of cognitive modes may well explain a variety of results in cognitive research and thus reduce the current isolation of research approaches. It may well turn out that, like the engineers in our study, we know more about our field than we have previously been able to express.

Reference Notes

1. Hammond, K. R. Unification of theory and research in judgment and

decision making (Center for Research on Judgment and Policy).

Unpublished manuscript, University of Coloradc, 1982.

References

Adelman, L. The influence of formal, substantive, and contextual task properties on the relative effectiveness of different forms of feedback in multiple-cue probability learning tasks. Organizational Behavior and Human Performance, 1981, 27, 423-442.

Anderson, B., Deane, D., Hammond, K. R., McClelland, G. H., & Shanteau, J. Concepts in judgment and decision research: Definitions, sources, interrelations, comments. New York: Praeger, 1981.

Brehmer, B. Policy conflict as a function of policy similarity and policy complexity. Scandinavian Journal of Psychology, 1972, 13, 208-221.

Brehmer, B. Social judgment theory and the analysis of interpersonal conflict. Psychological Bulletin, 1976, 83, 985-1003.

Brehmer, B. Preliminaries to a psychology of inference. Scandinavian Journal of Psychology, 1979, 21. 209-221.

Brehmer, B., & Hammond, K. R. Cognitive factors in interpersonal conflict. In D. Druckman (Ed.), Negotiations: Social-psychological perspectives. Beverly Hills: Sage, 1977.

Brooks, L. Nonanalytic concept formation and memory for instances. In E. Rosch and B. Lloyd (Eds.), Cognition and categorization. New York: Erlbaum, 1978.

Brunswik, E. The conceptual framework of psychology. In International
    encyclopedia of unified science (Vol. 1, No. 10). Chicago: University
    of Chicago Press, 1952.

Brunswik, E. Perception and the representative design of psychological
    experiments (2nd ed.). Berkeley: University of California Press, 1956.

Cohen, L. J. Can human irrationality be experimentally demonstrated? The
    Behavioral and Brain Sciences, 1981, 4(3), 317-370.

Duda, R., & Shortliffe, E. Expert systems research. Science, 1983, 220,
    261-268.

Edwards, W. Bayesian and regression models of human information
    processing--a myopic perspective. Organizational Behavior and Human
    Performance, 1971, 6, 639-648.

Einhorn, H. J., & Hogarth, R. M. Behavioral decision theory: Processes of
    judgment and choice. Annual Review of Psychology, 1981, 32, 53-88.

Einhorn, H. J., & Hogarth, R. M. Reply to commentaries. In G. Ungson & D.
    Braunstein (Eds.), Decision making: An interdisciplinary inquiry.
    Boston: Wadsworth, 1982.

Fisch, H. U., Hammond, K. R., & Joyce, C. R. B. On evaluating the severity
    of depression: An experimental study of psychiatrists. British Journal
    of Psychiatry, 1982, 140, 378-383.

Fox, J. Making decisions under the influence of memory. _Psychological Review_, 1980, _81_, 190-211.

Gillis, J. S. Interpersonal conflict: Effects of chlorpromazine and thiothixene on acute schizophrenic patients. In K. R. Hammond & C. R. B. Joyce (Eds.), _Psychoactive drugs and social judgment: Theory and research_. New York: Wiley, 1975.

Gillis, J. S. Sources of judgmental impairment in paranoid and non-paranoid schizophrenics. _Journal of Abnormal Psychology_, 1978, _87_, 587-596.

Hammond, K. R. Representative versus systematic design in clinical psychology. _Psychological Bulletin_, 1954, _51_, 150-159.

Hammond, K. R. Probabilistic functioning and the clinical method. _Psychological Review_, 1955, _62_(4), 255-262.

Hammond, K. R. (Ed.). _The psychology of Egon Brunswik_. New York: Holt, Rinehart, & Winston, 1966.

Hammond, K. R., & Brehmer, B. Quasi-rationality and distrust: Implications for international conflict. In L. Rappoport & D. Summers (Eds.), _Human judgment and social interaction_. New York: Holt, Rinehart, and Winston, 1973.

Hammond, K. R., McClelland, G. H., & Mumpower, J. _Human judgment and decision making: Theories, methods, and procedures_. New York: Praeger, 1980.

Hammond, K. R., Stewart, T. R., Brehmer, B., & Steinmann, D. O. Social

 judgment theory. In M. F. Kaplan & S. Schwartz (Eds.), Human judgment

 and decision processes. New York: Academic Press, 1975.

Hammond, K. R., & Summers, D. A. Cognitive control. Psychological Review,

 1972, 79, 58-67.

Hammond, K. R., & Wascoe, N. E. (Eds.). New directions for methodology of

 social and behavioral science: Realizations of Brunswik's representative

 design. San Francisco: Jossey-Bass, 1980.

Hastie, R. Social inference. Annual Review of Psychology, 1983, 34,

 511-542.

Holzworth, J. Intervention in a cognitive conflict. Organizational Behavior

 and Human Performance, in press.

Johnson-Laird, P., & Wason, P. A theoretical analysis of insight into a

 reasoning task. In P. Johnson & P. Wason (Eds.), Thinking: Readings in

 cognitive science. Cambridge, England: Cambridge University Press,

 1977.

Jones, C., & Harris, P. Insight into the law of large numbers: A comparison

 of Piagetian and judgment theory. Quarterly Journal of Experimental

 Psychology, 1982, 34a, 479-488.

Kahneman, D., Slovic, P., & Tversky, A. Judgments under uncertainty:

 Heuristics and biases. New York: Cambridge University Press, 1982.

Kahneman, D., & Tversky, A. On the study of statistical intuitions.

Cognition, 1982, 11, 123-141.

Kirwan, J., Chaput de Saintonge, D., Joyce, C. R. B., & Currey, H. Clinical

judgment analysis: Practical application in rheumatoid arthritis.

British Journal of Rhematology, in press.

Kyburg, H. The logical foundations of statistical inference. Dordrecht,

Holland: D. Reidel, 1974.

Kyburg, H., & Smokler, H. Studies in subjective probability (2nd ed.). New

York: Wiley, 1980.

Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. Expert and novice

performance in solving physics problems. Science, 1980, 208(4450),

1335-1342.

Levi, I. The enterprise of knowledge. Cambridge, MA: Massachusetts

Institute of Technology Press, 1980.

Lichtenstein, S., Fischhoff, B., & Phillips, L. Calibration of

probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic

& A. Tversky (Eds.), Judgment under uncertainty: Heuristics and biases.

Cambridge, England: Cambridge University Press, 1982.

Luce, R. D. Individual choice behavior. New York: Wiley, 1959.

Oskamp, S. Overconfidence in case-study judgments. Journal of Consulting

Psychology, 1965, 29, 261-265.

Scribner, S. Modes of thinking and ways of speaking: Culture and logic

    reconsidered. In P. Johnson-Laird & P. Wason (Eds.), Thinking: Readings

    in cognitive science. Cambridge, MA: Cambridge University Press, 1977.

Shanteau, J. An information-integration analysis of risky decision making.

    In M. F. Kaplan & S. Schwartz (Eds.), Human judgment and decision

    processes. New York: Academic Press, 1975.

Slovic, P., Fischhoff, B., & Lichtenstein, S. Behavioral decision theory.

    Annual Review of Psychology, 1977, 28, 1-39.

Slovic, P., & Lichtenstein, S. Comparison of Bayesian and regression

    approaches to the study of information processing judgment.

    Organizational Behavior and Human Performance, 1971, 6, 649-744.

Tversky, A. Elimination by aspects: A theory of choice. Psychological

    Review, 1972, 79, 281-299.

Tweney, R., Doherty, M., & Mynert, C. (Eds.). On scientific thinking. New

    York: Columbia University Press, 1981.

Wason, P., & Johnson-Laird, P. Psychology of reasoning: Structure and

    content. Cambridge, MA: Harvard University Press, 1972.

APPENDIX A

Characteristics of Engineers' Formulas for

Safety, Analytical Mode

Eleven engineers developed formulas that involved only additive combinations of the cues (median r_a = .467). Four of these involved only linear functions of the cues (median r_a = .281). For example:

$$Accident\ Rate = 12.8(23 - (SWIDTH + LWIDTH))/15 +$$
$$12.8(OBSPM + CURVEPM + INTPM)/29 +$$
$$6.4(TRAFMIX + PCTNPZ)/115$$

One expressed nonlinear functions with graphs (r_a = .518); six used tables to express nonlinear functions of cues (median r_a = .496).

Another engineer used only a multiplicative combination of variables (r_a = .670):

$$Accident\ Rate = .25(F_{SW} * F_{LW} * F_{NPZ} * F_{GRD} * F_{IPM} *$$
$$F_{OPM} * F_{ASL} * F_{TM} * F_{TV})$$

where each F is a one-dimensional table expressing a linear or nonlinear function of the cue.

The remaining nine engineers produced formulas that involved a hierarchy of organizing principles (median r_a = .436). For example, one configurally combined a subset of factors in a table and then added the result together with other factors. Five of these hierarchical formulas used adding and multiplying (median r_a = .436). For example:

$$Accident\ Rate = CURVEPM + (60 - AVESL)/10 -$$
$$(LWIDTH + SWIDTH - 23)/2 +$$
$$(TRAFVOL/10000)(INTPM + OBSPM - 3) +$$
$$(TRAFMIX * GRADE * PCTNPZ)/10000$$

Three engineers used adding as well as tables (median r_a = .390). One engineer used multiplying and tables (r_a = .641).

APPENDIX B

Maximal Guidance Procedure

This procedure was designed to give the engineer a maximum of guidance in constructing his formula for safety in order to prevent the commission of minor errors and to ensure that his formula adhered to principles of measurement theory with which he might not be familiar. The procedure consisted of the following steps:

-- Specify the answer scale.

-- Specify the scale for each input dimension and its overall relation to the answer scale, and identify possible interactions with other dimensions.

-- Group the input dimensions according to their redundancy, similarity, or mutual interactions.

-- Express the formula as a hierarchy of groups of variables.

-- Determine what organizing principle should be used at each level of hierarchy.

-- Specify the function form governing each dimension's input to its organizing principle.

-- Combine all the above information into one formula.

The engineer was guided through these steps by a series of forms which
contained instructions for the steps and choice points, and on which
intermediate steps were recorded. Two examples follow. The engineer also
received detailed tutorials about interactions and organizing principles as
part of the maximal guidance procedure.

Form 1: Answer Dimension Form

Name _____   Task _____

Answer dimension's units _____.

Range of possible answers: Low _____   High _____.

A "natural 0" on a scale means that when it is called "0" there really is
NONE of the quality being measured. If you had a natural 0, then it would
make sense to say that an "8" is twice as much of a thing as a "4"; but if
the 0 was arbitrary, it wouldn't have that sort of meaning.

For example, if you have savings of $10,000, you have twice as much money
as if you had $5000, because 10000 is twice 5000. Here the $0 is a natural 0.
But 32 degrees F is not twice as warm as 16 degrees F, because the 0 on the
temperature scale is picked arbitrarily. In other words, it does not have a
natural 0.

Does the answer dimension have a natural 0? Yes _____ No _____.

It is useful, when considering numbers that measure a dimension, to ask
whether the intervals between the numbers have consistent meaning, or whether
the numbers simply express order. For example, is the difference between a 1
and a 2 the same as the difference between an 11 and a 12? In the above
measures of money or temperature, the intervals do have consistent meaning.
However if we were to assign numbers to grades on a test, where A = 1, B = 2,
C = 3, D = 4, E = 5, and F = 6, the interval between 2 and 4 would be
different in meaning from the interval between 4 and 6. All the numbers
convey is that A is better than B, etc.

Do the intervals in the answer scale have a consistent meaning?
Yes _____ No _____.

Form 3: Choice of Organizing Principle

This form is for use in deciding what organizing principle to use for producing either the final answer or an intermediate product to be plugged in at a higher level in the hierarchy.

<u>Output.</u>

Is this the top level, producing the final answer? Yes _____ No _____.

If so, what are the units of the final answer? _____.

What is its range? Low _____ High _____.

If this is not the top level, then the output of this organizing principle will be input for an organizing principle at a higher level.

What organizing principle is used at the next higher level? _____.

What kind of input does it require? Units _____.

Range: Lowest point _____, Highest point _____;

Does it need to have a natural 0? _____.

<u>Input.</u>

List the input dimensions:

<u>Organizing Principle.</u>

What organizing principle do you want to use here? (Refer to Sheet 2 for guidance in your choice, and to the Forms 2-i and 2-g that you used to describe these dimensions to see what kinds of interaction they have with each other.)

Check one: Averaging ___. Multiplying ___. Table (Configural) ___.
           Other ___.

APPENDIX C

Independent Confirmation of the Quasi-rational
Properties of the Safety-Judgment Task

Evidence that the safety-judgment task was not fully susceptible to analytical treatment can be seen in the following examples of spontaneous comments by engineers during their attempts to construct a formula:

General.

"You kind of have to be able to juggle them all in your head, you know, in order to assign a weight to them."

"I kind of understand what I did."

"This is too hard! Can I form a committee?"

"I haven't the vaguest idea how this is going to come out."

Subject attempts to include curves and passing versus lane width.

"On the one hand, the lane width and shoulder width lends itself to giving the drivers a sense of security -pause- perhaps a sense of comfort on the road -pause- where the geometry is -pause- not like that. Oh boy, I don't know what to do with those."

Subject attempts to use average speed limit. (1)

"But whether or not you would have more accidents per mile typically
on something that is 30 miles an hour or 60 miles an hour is a little bit
ambiguous, and I don't know that I could really properly consider that.
My inclination is to leave it out for the time being. Either you leave it
out or conversely assume that it is the only valid measure of safety. I
think for the time being I'm going to leave it out."

Subject attempts to use average speed limit. (2)

"It's really difficult for me to get a feel for how that influences
highway safety. So rather than trying to guess, I just left it off,
recognizing that, you know, in a further generation or a further
refinement of the formula, you might be able to come to some conclusion
about the effect and include it."

APPENDIX D

Notes on Think Aloud Transcripts

Engineer #1

-organizing principle--has feel for multiplicative relationship but in
end he merely rank orders cues and assigns weights--additive

-%NPZ--he finds this cue ambiguous, but finally resolves it as
diagnostic--he calls it "inverted", "reciprocal" (implying its function
form), but then in his formula he uses "100 - %NPZ" (complement)--he
does not see this cue as that important

-ASL--also seen as ambiguous--leans toward using it predictively, but
then leaves it out of the formula because he is unsure how to use it--he
mentions that, with regard to ASL, what is more dangerous is a high
differential, and a low posted speed limit may increase this (some cars
obey the low SL while other cars disregard what they may see as
unnecessarily low)

[not enough information for engineer to know the characteristics of a
particular highway regarding their influence on ASL; also, he is
unaware of how to express such a relationship in a generalized formula;
therefore, cue is simply left out of formula]

-indicates possibilities for a more complex organizing principle (than
additive), i.e., nonlinear - TV2 - "but need data to make that judgment"

Engineer #5

-Engineer #5 initially breaks accidents down into the subcategories of
property damage, injury, and fatality--elements affect different types
of accidents; these can combine differently to get the same accident
rate per mile

-mentions cues that relate to each other - "a lot of these things are
redundant", e.g., OPM and SW [cue interactions]

others: %NPZ/SPD/TM, GRD/TM, CPM/LW --"factors can aggravate each
other"

-mentions that cues may be "cumulative" rather than additive, i.e.,
interactive

-the relationship of TM to accidents is geometric (logarithmic?)--"0-5%,
6-10%, 11-20%, 21-35%" each expresses a range of ratings that are "about
the same"

-he aims for a norm/baseline as he develops his formula and attempts to produce results within the desired range. When the results are slightly off, he goes back to one of his matrices and adjusts the "range" for SW [actually he has adjusted the beta weight of this cue in his equation]

Interview:

-how would he use historical research (which he mentions he would like to consult if he could)--in the past he has "taken an average rate and modified it based on whether you provide certain features or detract from those features"; he would look at how "features stack up and influence rates"

-he indicates that he knows he would use a regression equation if he had the data

-he also mentions that if the data were available, the initial three categories (property damage, injury, fatality) would be kept separate--e.g., some barriers may reduce fatalities yet increase property damage. Therefore, each individual equation should be different for the three categories--of these, the property damage formula is the simplest, but the other two are not as high(?) as property damage, and may be insignificant on the same scale (especially fatalities), causing the formula to be basically a property damage formula (which, again, is simpler)

-he indicates a definite awareness of multiplicative relationships

-confidence in method--"I would hesitate to create an equation for all roadway conditions because you just have to take a look at where the roadway is--"

-capacity was easier for him--was well researched, time-space relationship rather than individual characteristics that could influence (because of variance?)

Engineer #8

-also starts with optimum value of each characteristic

-%NPZ--"skrewed" relationship--"can't get a handle on how to get that incorporated into this (needs to use a matrix?)

-uses trial and error to obtain the relationship (function) which "looks" right (intuitive feel for relationship?)

-converts all ranges to a 1-10 scale based on an optimum which may not be an end point

-SPD--60 set as optimum irregardless of other factors [no interactions]

-OPM--range (0-18) is too complicated to calibrate so "satisfices" by
spreading 1-10 ratings out along the range;   therefore some cue values
have the same rating (arbitrary?)--i.e., 12 and 13 both equal a rating
of 10.

-uses a weighted average organizing principle

-he thought safety "was going to be easy because it would fit into one of
the slots like I did before, but it doesn't." (He had done both
aesthetics and capacity)--he discovers that he cannot just subtract off
the number of cars for each limiting factor like in capacity

-he observes that this is like most any formula--"it gives you some kind
of relativity of the things."

-with regard to the memory task, he says there is "no way in the world"
he could remember his weighting system (how analytic was he? or does
this generalization hold?)

Engineer #10

-prioritizes cues first and then arranges them in a linear additive
manner, but not weighted average. -originally ranks cues by class
(cross section, geometry, traffic) and then give subrankings, but then
in the weighting process each cue is examined independently and the
overall rankings are altered [no cue interactions]

-realizes he has to normalize cue ranges, i.e., give ratings to each
possible cue value on a scale from 1-10

-he forgets that his answer should fall between 0-32--initially comes up
with a range in the hundreds, but he is unsure of the endpoints--he says
one could work them out

-After trying out his formula on three examples, he realized that he had
framed the wrong relationships, i.e, most of the cues had a negative
relation (correlation) with safety rather than a positive one. He
compensated for this by switching the scale to low=safe/high=unsafe and
then readjusting the two positively related cues to reflect this shift.
All of his cues were used predictively except for %NPZ, which he regards
as negatively related to safety [and could be diagnostic or
predictive--%NPZ could indicate an unsafe road or it could frustrate
drivers and cause more accidents (both negative)]]

-why low confidence? (5)--"not enough thought" put into it--would like
to test over a range of highways "knowing the data".

-memory task--He has problems remembering all the cues, much less the
numbers he assigned...

Engineer #12

   -begins to combine cues even as he reads their definitions and ranges,
    e.g., %NPZ/TV/TM

   -approaches safety rating from a "relative" standpoint since the chances
    of something being zero [are] slim" [answer for safety is on a relative
    scale though cues are presented with their absolute ranges for the
    sample--example of prior knowledge interferring with an analytic
    approach?]

   -changes cues to relative scales also--i.e., (LW+SW) ranges from -2 to +2

   -wants an interaction between %NPZ and TM but their multiplication can
    produce a zero if one term is a zero
    wants to say, that if both %NPZ and TM are high, there is an interaction
    [nonlinear]

   -mentions "duplication" (compare Engineer #5's complaint about
    "redundancy")--have no passing on steep grades and curves, "so it's all
    kind of melded in together and somehow we're duplicating each other"

   -"intuitively" knows the kind of relationships he wants to express, i.e.,
    what kind of an effect a certain cue interaction should have on the
    overall safety rating, but cannot express it--goes through trial and
    error

   -he winds up with a variable in the denominator that could be zero--he
    responds by saying "I won't allow it to be zero".

   -he constantly tries out the formula he has so far--to make sure it
    sounds right and nothing is too far off (dealing with more concrete
    examples)

   -he wants his worse case to be 32, so in weighting he thinks in
    approximate percentages of 32

   -SPDL--diagnostic--"the lower the speed limit is, the worse the situation
    is."

   -ultimately, his strategy is to weight each "group", i.e., he produces a
    "traffic number", by normalizing the ranges of involved cues so that
    they add up to a percentage of 32.
    [this comes out linear in spite of his earlier expression as to the
    nonlinearity--he tries to show some relationship by grouping in brackets
    and multiplying or diving by a common factor]

-SPDL--not weighted as heavily because "in setting the speed limit, it
considered all of the other things that were taken into consideration"
[diagnostic; but he does not see it as an interaction here between SPDL
and other cues]

-during the discussion he expresses an interaction between LW and OPM
that he did not include in his formula

-some cues can be plugged directly into his formula (e.g., IPM), while
others (e.g., SW) have been converted to a different scale--negative,
neutral, or positive [SW, as he expresses it verbally, may be
third-order relationship]

Engineer #15

-he recognizes that the safety rating must be in terms of million vehicle
miles travelled and so TV will be a component of the rating rather than
strictly a factor

-wants to develop weights that would predict the number of accidents
which would then be correlated with the required range (0-32) [he asks
"can I do that?"]

-uses TV on both sides of his equation (now using as a factor)

-he sees the complexity of the task but attempts to "keep it simple"

-weighting factors were the "factors of importance", which then had to be
put in the same "context" (normalized) by giving them a safety
rating--produced a "relative safety value" (actually he just changed the
range on each cue to 1-5) but relationships are not quite linear

-%NPZ--diagnostic--the higher the percentage, the worse the condition

-SPD--predictive--the lower the speed limit, the safer

-multiplies ranges by weighting factors to amplify them

-assumes a linear relationship between cues because he has "built any
curve aspects into the judgments here" (the cue-criterion relationships)

-he comments that this empirical formula is probably too
complicated--after trying it out, one would probably find that he could
get just as good results by dealing with four or five factors rather
than all of the dimensions. He nevertheless used all of them in his
formula "because there was a judgment built into this whole thing in the
first place that these factors were important."

-places TV in the denominator to give "a means of comparing different
length segments" and "differing volumes"--what he doesn't realize is
that the rating is already in terms of million vehicle miles travelled
(or maybe he just forgets this)

-in going over his steps he recalls wanting to "use a system where I come
up with an intuitive judgment about the safety of the thing as judged by
the dimensions given" [underlining ours]

-"if you're going to have a practical formula, you've got to have a way
where people can use it quickly." [In the process of developing a
"practical formula" that can be applied easily does one oversimplify
complex relationships--thereby reducing validity? (and accuracy?)]

APPENDIX E

Comparison of Repeated Trials Reliability with Cognitive Control

Fourteen of the engineers had higher R_c's than R_s's in the intuitive

mode. This finding suggests that these engineers were using information in a

nonlinear way, and/or that the pictorial presentation contained information

not present in the ten cues provided in the bar-graph and formula

presentations. Both possibilities were investigated, and neither can be ruled

out with certainty. The evidence suggests, however, that (a) the engineers

did not make systematic or substantial use of information in a nonlinear

fashion; (b) in the film-strip condition they occasionally used information

that was not present in the other tasks; (c) the engineers' occasional use of

additional information in the intuitive task does not account for the higher

repeated trials reliabilities.

The contribution of the nonlinear component of the Lens Model Equation

was investigated to determine whether a significant relation existed between

the portions of the variances of the criterion and the judgments that were not

explained by their respective linear models. Minimal evidence of such

nonlinear knowledge was found. The mean absolute value of the nonlinear

component ($C \sqrt{1 - R_e^2} \sqrt{1 - R_s^2}$) was .056 in the intuitive mode and .048 in

the quasi-rational mode.

That some engineers in the intuitive condition made use of information
present only in the film strips, at least on some trials, is almost certainly
true. For instance, residential driveways entering the highway were evident
and could have been taken into account.

The most plausible explanation for the higher repeated trials reliability
in the film-strip condition than in the bar-graph condition is that the
engineers recognized the repeated film strips and recalled their previous
answers. There are two reasons for accepting this explanation: (a) it was
obvious to the engineers that they were viewing a film for a second time, and
some of them remarked that they remembered their previous answer; (b) whereas
repeated trials reliability was higher than multiple R in the film-strip
condition, it was higher for only two engineers in the bar graph condition;
and no engineer in this condition indicated that he recognized a repeated
trial. In short, the fact that fourteen of the 21 engineers had higher
repeated trials reliabilities in the film-strip condition is almost certainly
the result of an artifact of memory.

APPENDIX F

Confidence Questions

Answer Confidence. In the analytical mode the engineer was asked, "How confident are you that your formula will work well with any particular 2-lane rural Colorado highway?" once after producing the formula and again at the end of the session. The mean of his responses is the measure of his answer confidence in the analytical mode.

Method Confidence. In the intuitive and quasi-rational modes the mean of the confidence ratings on the following three questions from the Self Report form is the measure of method confidence:

1. How confident are you that your method for making these safety judgments is correct?

2. How well does this presentation help you make safety judgments?

3. How accurate do you think one could be when making these judgments in this way?

The measure for analytical method confidence is the mean of the ratings on the following questions:

1. How confident are you that your formula is correct? (answered at two different times)

2. How accurate do you think a formula for judging the capacity of highways can be?

OFFICE OF NAVAL RESEARCH

Engineering Psychology Group

TECHNICAL REPORTS DISTRIBUTION LIST

OSD

CAPT Paul R. Chatelier
Office of the Deputy Under Secretary
  of Defense
OUSDRE (E&LS)
Pentagon, Room 3D129
Washington, D. C.   20301

Dr. Dennis Leedom
Office of the Deputy Under Secretary
  of Defense (C^{3}I)
Pentagon
Washington, D. C.   20301

Department of the Navy

Engineering Psychology Group
Office of Naval Research
Code 442 EP
Arlington, VA   22217 (2 cys.)

Aviation & Aerospace Technology
  Programs
Code 210
Office of Naval Research
800 North Quincy Street
Arlington, VA   22217

Communication & Computer Technology
  Programs
Code 240
Office of Naval Research
800 North Quincy Street
Arlington, VA   22217

Physiology & Neuro Biology Programs
Code 441NB
Office of Naval Research
800 North Quincy Street
Arlington, VA   22217

Department of the Navy

Tactical Development & Evaluation
  Support Programs
Code 230
Office of Naval Research
800 North Quincy Street
Arlington, VA   22217

Manpower, Personnel & Training
  Programs
Code 270
Office of Naval Research
800 North Quincy Street
Arlington, VA   22217

Mathematics Group
Code 411-MA
Office of Naval Research
800 North Quincy Street
Arlington, VA   22217

Statistics and Probability Group
Code 411-S&P
Office of Naval Research
800 North Quincy Street
Arlington, VA   22217

Information Sciences Division
Code 433
Office of Naval Research
800 North Quincy Street
Arlington, VA   2217

CDR K. Hull
Code 230B
Office of Naval Research
800 North Quincy Street
Arlington, VA   22217

Department of the Navy

Special Assistant for Marine Corps
  Matters
Code 100M
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Dr. J. Lester
ONR Detachment
495 Summer Street
Boston, MA 02210

Mr. R. Lawson
ONR Detachment
1030 East Green Street
Pasadena, CA 91106

CDR James Offutt, Officer-in-Charge
ONR Detachment
1030 East Green Street
Pasadena, CA 91106

Director
Naval Research Laboratory
Technical Information Division
Code 2627
Washington, D. C. 20375

Dr. Michael Melich
Communications Sciences Division
Code 7500
Naval Research Laboratory
Washington, D. C. 20375

Dr. J. S. Lawson
Naval Electronic Systems Command
NELEX-06T
Washington, D. C. 20360

Dr. Robert E. Conley
Office of Chief of Naval Operations
Command and Control
OP-094H
Washington, D. C. 20350

CDR Thomas Berghage
Naval Health Research Center
San Diego, CA 92152

Department of the Navy

Dr. Robert G. Smith
Office of the Chief of Naval
  Operations, OP987H
Personnel Logistics Plans
Washington, D. C. 20350

Dr. Andrew Rechnitzer
Office of the Chief of Naval
  Operations, OP 952F
Naval Oceanography Division
Washington, D. C. 20350

Combat Control Systems Department
Code 35
Naval Underwater Systems Center
Newport, RI 02840

Human Factors Department
Code N-71
Naval Training Equipment Center
Orlando, FL 32813

Dr. Alfred F. Smode
Training Analysis and Evaluation
  Group
Orlando, FL 32813

CDR Norman E. Lane
Code N-7A
Naval Training Equipment Center
Orlando, FL 32813

Dr. Gary Poock
Operations Research Department
Naval Postgraduate School
Monterey, CA 93940

Dean of Research Administration
Naval Postgraduate School
Monterey, CA 93940

Mr. H. Talkington
Ocean Engineering Department
Naval Ocean Systems Center
San Diego, CA 92152

Department of the Navy

Mr. Paul Heckman
Naval Ocean Systems Center
San Diego, CA 92152

Dr. Ross Pepper
Naval Ocean Systems Center
Hawaii Laboratory
P. O. Box 997
Kailua, HI 96734

Dr. A. L. Slafkosky
Scientific Advisor
Commandant of the Marine Corps
Code RD-1
Washington, D. C. 20380

Dr. L. Chmura
Naval Research Laboratory
Code 7592
Computer Sciences & Systems
Washington, D. C. 20375

HQS, U. S. Marine Corps
ATTN: CCA40 (Major Pennell)
Washington, D. C. 20380

Commanding Officer
MCTSSA
Marine Corps Base
Camp Pendleton, CA 92055

Chief, C^3 Division
Development Center
MCDEC
Quantico, VA 22134

Human Factors Technology Administrator
Office of Naval Technology
Code MAT 0722
800 N. Quincy Street
Arlington, VA 22217

Commander
Naval Air Systems Command
Human Factors Programs
NAVAIR 334A
Washington, D. C. 20361

Department of the Navy

Commander
Naval Air Systems Command
Crew Station Design
NAVAIR 5313
Washington, D. C. 20361

Mr. Philip Andrews
Naval Sea Systems Command
NAVSEA 03416
Washington, D. C. 20362

Commander
Naval Electronics Systems Command
Human Factors Engineering Branch
Code 81323
Washington, D. C. 20360

Larry Olmstead
Naval Surface Weapons Center
NSWC/DL
Code N-32
Dahlgren, VA 22448

Mr. Milon Essoglou
Naval Facilities Engineering Command
R&D Plans and Programs
Code 03T
Hoffman Building II
Alexandria, VA 22332

Captain Robert Biersner
Naval Medical R&D Command
Code 44
Naval Medical Center
Bethesda, MD 20014

Dr. Arthur Bachrach
Behavioral Sciences Department
Naval Medical Research Institute
Bethesda, MD 20014

Dr. George Moeller
Human Factors Engineering Branch
Submarine Medical Research Lab
Naval Submarine Base
Groton, CT 06340

Department of the Navy

Dr. Edgar M. Johnson
Technical Director
U. S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Director, Organizations and
  Systems Research Laboratory
U. S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Technical Director
U. S. Army Human Engineering Labs
Aberdeen Proving Ground, MD 21005

Department of the Air Force

U. S. Air Force Office of Scientific
  Research
Life Sciences Directorate, NL
Bolling Air Force Base
Washington, D. C. 20332

AFHRL/LRS TDC
Attn: Susan Ewing
Wright-Patterson AFB, OH 45433

Chief, Systems Engineering Branch
Human Engineering Division
USAF AMRL/HES
Wright-Patterson AFB, OH 45433

Dr. Earl Alluisi
Chief Scientist
AFHRL/CCN
Brooks Air Force Base, TX 78235

Foreign Addressees

Dr. Daniel Kahneman
University of British Columbia
Department of Psychology
Vancouver, BC V6T 1W5
Canada

Foreign Addressees

Dr. Kenneth Gardner
Applied Psychology Unit
Admiralty Marine Technology
  Establishment
Teddington, Middlesex TW11 OLN
England

Director, Human Factors Wing
Defence & Civil Institute of
  Environmental Medicine
Post Office Box 2000
Downsview, Ontario M3M 3B9
Canada

Dr. A. D. Baddeley
Director, Applied Psychology Unit
Medical Research Council
15 Chaucer Road
Cambridge, CB2 2EF England

Other Government Agencies

Defense Technical Information Center
Cameron Station, Bldg. 5
Alexandria, VA 22314 (12 copies)

Dr. Craig Fields
Director, System Sciences Office
Defense Advanced Research Projects
  Agency
1400 Wilson Blvd.
Arlington, VA 22209

Dr. M. Montemerlo
Human Factors & Simulation
Technology, RTE-6
NASA HQS
Washington, D. C. 20546

Dr. J. Miller
Florida Institute of Oceanography
University of South Florida
St. Petersburg, FL 33701

Other Organizations

Dr. Robert R. Mackie
Human Factors Research Division
Canyon Research Group
5775 Dawson Avenue
Goleta, CA 93017

Dr. Amos Tversky
Department of Psychology
Stanford University
Stanford, CA 94305

Dr. H. McI. Parsons
Human Resources Research Office
300 N. Washington Street
Alexandria, VA 22314

Dr. Jesse Orlansky
Institute for Defense Analyses
1801 N. Beauregard Street
Alexandria, VA 22311

Professor Howard Raiffa
Graduate School of Business
  Administration
Harvard University
Boston, MA 02163

Dr. T. B. Sheridan
Department of Mechanical Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139

Dr. Arthur I. Siegel
Applied Psychological Services, Inc.
404 East Lancaster Street
Wayne, PA 19087

Dr. Paul Slovic
Decision Research
1201 Oak Street
Eugene, OR 97401

Dr. Harry Snyder
Department of Industrial Engineering
Virginia Polytechnic Institute and
  State University
Blacksburg, VA 24061

Other Organizations

Dr. Ralph Dusek
Administrative Officer
Scientific Affairs Office
American Psychological Association
1200 17th Street, N. W.
Washington, D. C. 20036

Dr. Robert T. Hennessy
NAS - National Research Council (COHF)
2101 Constitution Avenue, N. W.
Washington, D. C. 20418

Dr. Amos Freedy
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

Dr. Robert C. Williges
Department of Industrial Engineering
  and OR
Virginia Polytechnic Institute and
  State University
130 Whittemore Hall
Blacksburg, VA 24061

Dr. Meredith P. Crawford
American Psychological Association
Office of Educational Affairs
1200 17th Street, N. W.
Washington, D. C. 20036

Dr. Deborah Boehm-Davis
General Electric Company
Information Systems Programs
1755 Jefferson Davis Highway
Arlington, VA 22202

Dr. Ward Edwards
Director, Social Science Research
  Institute
University of Southern California
Los Angeles, CA 90007

Dr. Robert Fox
Department of Psychology
Vanderbilt University
Nashville, TN 37240

Dr. Charles Gettys
Department of Psychology
University of Oklahoma
455 West Lindsey
Norman, OK 73069

Dr. Kenneth Hammond
Institute of Behavioral Science
University of Colorado
Boulder, CO 80309

Dr. James H. Howard, Jr.
Department of Psychology
Catholic University
Washington, D. C. 20064

Dr. William Howell
Department of Psychology
Rice University
Houston, TX 77001

Dr. Christopher Wickens
Department of Psychology
University of Illinois
Urbana, IL 61801

Mr. Edward M. Connelly
Performance Measurement
  Associates, Inc.
410 Pine Street, S. E.
Suite 300
Vienna, VA 22180

Professor Michael Athans
Room 35-406
Massachusetts Institute of
  Technology
Cambridge, MA 02139

Dr. Edward R. Jones
Chief, Human Factors Engineering
McDonnell-Douglas Astronautics Co.
St. Louis Division
Box 516
St. Louis, MO 63166

Dr. Babur M. Pulat
Department of Industrial Engineering
North Carolina A&T State University
Greensboro, NC 27411

Dr. Lola Lopes
Information Sciences Division
Department of Psychology
University of Wisconsin
Madison, WI 53706

Dr. A. K. Bejczy
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91125

Dr. Stanley N. Roscoe
New Mexico State University
Box 5095
Las Cruces, NM 88003

Mr. Joseph G. Wohl
Alphatech, Inc.
3 New England Executive Park
Burlington, MA 01803

Dr. Marvin Cohen
Decision Science Consortium
Suite 721
7700 Leesburg Pike
Falls Church, VA 22043

Dr. Wayne Zachary
Analytics, Inc.
2500 Maryland Road
Willow Grove, PA 19090

Dr. William R. Uttal
Institute for Social Research
University of Michigan
Ann Arbor, MI 48109

Dr. William B. Rouse
School of Industrial and Systems
  Engineering
Georgia Institute of Technology
Atlanta, GA 30332

Other Organizations

Dr. Richard Pew
Bolt Beranek & Newman, Inc.
50 Moulton Street
Cambridge, MA 02238

Dr. Hillel Einhorn
Graduate School of Business
University of Chicago
1101 E. 58th Street
Chicago, IL 60637

Dr. Douglas Towne
University of Southern California
Behavioral Technology Laboratory
3716 S. Hope Street
Los Angeles, CA 90007

Dr. David J. Getty
Bolt Beranek & Newman, Inc.
50 Moulton street
Cambridge, MA 02238

Dr. John Payne
Graduate School of Business
  Administration
Duke University
Durham, NC 27706

Dr. Baruch Fischhoff
Decision Research
1201 Oak Street
Eugene, OR 97401

Dr. Andrew P. Sage
School of Engineering and
  Applied Science
University of Virginia
Charlottesville, VA 22901

Denise Benel
Essex Corporation
333 N. Fairfax Street
Alexandria, VA 22314

Psychological Documents (3 copies)
ATTN: Dr. J. G. Darley
N 565 Elliott Hall
University of Minnesota
Minneapolis, MN 55455