Stylistic Components of Human Judgment: The Generality of Individual Differences

Nerella V. Ramanaiah
Southern Illinois University at Carbondale
Lewis R. Goldberg
University of Oregon and Oregon Research Institute

In an effort to discover the generality of individual differences in judgmental processes, 12 potential styles were assessed from the predictions made by 86 judges across two or four replications of each of four judgmental tasks. The analyses focused on the internal consistency (across replications, within occasions), temporal stability (across occasions, within tasks), and intertask congruence of each of the judgmental variables. The findings suggested that five variables—linear predictability, judgmental consistency, subjective complexity, differentiation, and confidence—have sufficient generality to assume the provisional status of judgmental styles. Highly significant relationships were found, for male though not for female judges, between each of the provisional styles and a number of personological and demographic variables, including scales from the MMPI, CPI, EPPS, and SVIB.

The last decade has witnessed an explosion of psychological research in human judgment and decision-making. One focus of this broad scientific enterprise has been directed at the processes by which information is amalgamated (e.g., Anderson, 1968; Edwards, 1965; Goldberg, 1968; Hammond, Hursch, & Todd, 1964; Hoffman, 1960), in the hope of eventually uncovering any systematic biases endemic to intuitive judgments and predictions (e.g., Kahneman & Tversky, 1973; Tversky & Kahneman, 1973). Most of this research, which has recently been reviewed by Slovic and Lichtenstein (1973), has been focused on the adequacy of various mathematical models as representations of human judgments, choices, or other types of decisions.

For example, in studies carried out within the framework of Egon Brunswik's lens model of human judgment, a sizeable number of profiles of quantified values from a set of cues are presented sequentially to a subject or "judge," who is instructed to use the multivariate information from each profile in turn to make a judgment or prediction. These judgments are then employed as the dependent variable in a standard linear regression analysis, with the cue values used as the independent variables. By regressing the judgments onto the cues, a linear model of the judge is produced; the regression coefficients in the model are used to estimate the relative weights of each cue in this linear representation. The multiple correlation between the cue values and the judgments serves as an index of the extent to which the linear model "captures the policy" of that judge (Naylor & Wherry, 1965). Generally, one or more 'models is fitted to the judgments of an individual subject, and the adequacy of each such representation is then assessed.

Paradoxically, however, individual differences in these idiographic representations have rarely been explored. While judgmental investigators have typically observed substantial individual differences among judges in their subjective feelings of confidence in their judgments, in their accuracy, in the reliability of their judgments over time, in their linear predictability, in the relative weights they employ for each of the cues, and in the extent to which they are able to infer those weights when asked to do so, there have heretofore been no systematic investigations of the *generality* of these individual differences across occasions or across judgmental tasks. It is not uncommon to hear references to some highly "accurate," "confident," or "linear" judge, as if these individual differences were general dispositions. However, given the current jaundice toward personality traits (e.g., Mischel, 1968), it is now incumbent upon investigators of individual differences to demonstrate, rather than merely to assume, their generality.

As used here, the concept of generality refers to the internal consistency, temporal stability, and trans-situational congruence of an individual difference (Campbell & Fiske, 1959; Cronbach, Gleser, Nanda, & Rajaratnam, 1972). The terms "trait" or "style" have come to be used to refer to a variable with a high degree of such generality. Extrapolating from studies of person perception (Shrauger & Altrocchi, 1964), one could hazard the guess that individual differences in judgmental accuracy should have minimal cross-task generality (e.g., Crow & Hammond, 1957), while individual differences in confidence are likely to be of quite substantial generality (e.g., Johnson, 1957; Little, 1961). The present study provides a test of these extrapolations by examining the generality of individual differences on twelve judgmental variables across four multiple-cue tasks.

The first problem is to identify potential individual differences within the framework of a typical judgmental investigation. Subsequently, one should investigate the extent to which these individual differences are consistent over repeated administrati ns of the same judgmental task. For variables which show such temporal consistency, one can then discover the extent of their generality across different judgment tasks. And, finally, for those surviving variables whose consistency and generality has been established—and which therefore assume the provisional status of judgmental styles—one can investigate some of their broader personality correlates. These are the major goals of the present investigation.

Method

Judgment Tasks

Four tasks were employed. In each, the judges were provided a booklet of profiles containing quantified multivariate information and told that the profiles were constructed from information describing former University of Oregon freshmen. They were asked to approach each task as if they were members of a selection committee who were required to assess applicants on the particular attribute in question. Judges worked individually on one task per experimental session, without any time limit. They returned once a week over a two-month period, to complete these four tasks, plus others that required various psychophysical and aesthetic judgments. To enhance motivation, a sizeable cash prize was offered to the most accurate judge.

The first two tasks were selected on the basis of past judgmental research (e.g., Hoffman, 1960), in which they appeared to be quite involving for most judges and to stimulate a considerable range of individual differences. In one of these tasks, the judges predicted the intelligence of each of 120 college students, on the basis of the following nine cues: socioeconomic status, mother's education, extent of self-support, college credit hours attempted, high school grade point average, and scores from an English effectiveness test, a study habits inventory, an "emotional anxiety" scale, and a "responsibility" scale. In a second task, the judges predicted the "sociability" of another 120 college students, on the basis of their scores on eight

scales from the *Edwards Personal Preference Schedule:* Abasement, Affiliation, Change, Deference, Dominance, Exhibition, Heterosexuality, and Succorance. For each of these tasks, the 120 profiles comprised two replications of the same 60 profiles, which were given different identification numbers and then shuffled together. Criterion information was not available for either of these tasks, and consequently they can provide no information about individual differences in judgmental accuracy.

In contrast, however, criterion information was available for the remaining two tasks. These tasks required predictions of the freshman grade point average and of the extent of involvement in campus leadership activities, respectively, for each of 160 college students. The same set of nine cues was used for both tasks: *(a)* sex, *(b)* number of siblings, *(c)* father's education, *(d)* mother's education, *(e)* high school grade point average, *(f)* a college aptitude test score, *(g)* an English effectiveness test score, *(h)* the score from a study habits inventory, and *(i)* the number of hours per week that the student was employed. The set of 160 profiles was composed of two replications of 80, which were given different identification numbers and then intermingled. All judgments were made four times: the judges made their two sets of predictions for each task on each of two occasions, the second one about one month after the first. The two tasks differed considerably in their intrinsic difficulty; the linear predictability of the GPA criterion (i.e., the multiple correlation between the nine predictors and the actual criterion values) was approximately .80, while that for the leadership criterion was approximately .50.[1] Information concerning the four tasks is summarized in Table 1.

The profiles for each of the four tasks were administered in the same sequence to all judges.

[1]The intercorrelations among the predictors for all four tasks, and the validity coefficients for the GPA and leadership tasks, are available in Ramanaiah (1971).

Table 1

The Four Judgment Tasks

Prediction	Freshman GPA	Campus Leadership	Intelligence	Sociability
Number of cues	9	9	9	8
Number of unique profiles	80	80	60	60
Number of profiles per occasion	160	160	120	120
Number of occasions	2	2	1	1
Total number of replications	4	4	2	2
Criterion available	Yes	Yes	No	No
Maximum Predictability (R_e)	.80	.50	-	-
Estimated $\hat{R}$ (corrected for shrinkage)	.77	.39	-	-

Fifteen practice profiles were inserted at the beginning of each booklet, and separate answer sheets were used for recording responses. Ten-point rating scales were used to record all judgments. After making each judgment, the subjects were asked to indicate their confidence in that prediction by assigning a number between 0 and 1.00 to represent the probability of their being correct. Immediately following each task, the judges were asked to describe their prediction policies by distributing 100 points according to the relative importance of each cue in their overall judgments.

Judgmental Variables

In search of potential judgmental styles, the following variables were selected for preliminary analysis.

Measures of predictability and reliability:

Linear predictability. The linear predictability of the judge: the multiple correlation (R_s) between the cue values and the judgments for each replication of the task.

Consistency. The correlation between the two sets of judgments made for the two replications of each task on the same occasion.

Linear Consistency. The linear component of judgmental reliability: the correlation between the predicted values from a linear model of the judgments for one replication and those from a linear model of the judgments for the second replication.

Nonlinear Consistency. The nonlinear component of judgmental consistency: the correlation between the residual values of the judgments from both replications after the linear components of both sets of judgments have been removed.

Measures of accuracy and its components:

Accuracy. The validity coefficient of the judge: the correlation between the judgments and the actual criterion values.

Linear Accuracy. The linear component of judgmental accuracy: the correlation between the predicted values from a linear model of the judgments and those from a linear model of the criterion (see Tucker, 1964).

Nonlinear Accuracy. The nonlinear component of judgmental accuracy: the correlation between the residual values of the criterion and the residual values of the judgments after the linearly predictable components in both the criterion and the judgments have been removed (see Hammond, Hursch, & Todd, 1964; Goldberg, 1970).

Measures of complexity and insight:

Observed complexity. The amount of available information the judge appeared to have used: the standard deviation of the set of relative weights (Hoffman, 1960) associated with each of the cues, when the judgments are regressed on the cues.[2]

Subjective complexity. The amount of available information the judge felt he was using: the standard deviation of the subjective weights, based upon the judge's distribution of 100 points across the cues at the end of the task.

Insight. The correlation between a judge's subjective weights and the corresponding relative weights when the judgments are regressed on the cues.[3]

Measures of differentiation and confidence:

Differentiation. The standard deviation of the set of judgments for each replication of a task.

Confidence. The mean of the judge's ratings of confidence in each of his judgments for each replication of a task.

Analysis

The general strategy of analysis involved the following steps for each of the twelve judgmental variables:

a) Every judge's score was computed for each

---

[2]A number of alternative indices of observed complexity were also analyzed, including *(a)* the number of predictors with relative weights $\geq .10$, and the standard deviations of *(b)* the validity coefficients, *(c)* the regression coefficients, and *(d)* the standardized regression coefficients (beta weights).

[3]A number of alternative indices of "insight" were also analyzed, including the correlation between the subjective weights and *(a)* the validity coefficients, *(b)* the regression coefficients, and *(c)* the standardized regression coefficients (beta weights).

replication of each of the four judgment tasks. Whenever these values were correlations, their z-transformed equivalents were used.

b) The mean values of each variable were tested for possible sex differences. The analyses described below were carried out separately for the male and the female samples, as well as for the entire sample.

c) The scores described in step a) were inter-correlated across the judges. Variables which demonstrated substantial generality—as evidenced by high coefficients of internal consistency (across replications, within tasks and occasions), high coefficients of temporal stability (across occasions, within tasks), and statistically significant inter-task correlations—were considered as provisional judgmental styles.

d) For every such provisional style, each judge's scores were averaged over all replications, all occasions, and all tasks.

e) Intercorrelations among the styles were obtained across subjects, using the composite scores from step d).

f) The composite scores for the styles were then correlated with demographic variables and personality scale scores.

Subjects and Personality Variables

Eighty-six undergraduate students (37 men and 49 women) were paid a generous hourly rate to participate in the experiment. These subjects were a sample of the 95 male and 108 female students who had previously completed a course in introductory psychology at the University of Oregon, and who had taken the *Minnesota Multiphasic Personality Inventory (MMPI)*, the *California Psychological Inventory (CPI)*, the *Edwards Personal Preference Schedule (EPPS)*, and various other personality measures as a requirement for that course. A more complete description of the sample, and of the personality scales employed in this project, can be found in Goldberg and Jones (1969). At the end of the judgment sessions, the subjects completed the *Strong Vocational Interest Blank (SVIB)*, plus a brief biographical inventory.

Results

The means and standard deviations of each of the judgmental variables in the male and the female samples, for each replication of the four tasks, can be found in Ramanaiah (1971). Table 2 presents the mean values from the first replication of each task; these values are typical of those obtained from the other replications. The average values in the male and female samples were quite similar, and therefore those presented in Table 2 are based on the total sample of subjects.

The linear predictability values of the average judge, the mean multiple correlations between cues and judgments (R_s), were quite similar to those based on past research with these tasks, varying from .94 and .92 for predictions of intelligence and GPA to .83 and .75 for predictions of sociability and leadership. As is typical in tasks of this kind, the across-replications consistency of the average judge was quite high, though these correlations were higher for predictions of intelligence ($\bar{r} = .90$) and GPA ($\bar{r} = .87$) than for sociability ($\bar{r} = .73$) and leadership ($\bar{r} = .65$). The average linear consistency values were close to unity (.93 to .98), while the corresponding nonlinear averages were considerably lower (.24 to .29). Judgmental accuracy varied dramatically between the two tasks for which criterion information was available; the average judge was far superior in predicting GPA ($\bar{r} = .67$) than he was in predicting leadership ($\bar{r} = .13$). In this study, as in most others, accuracy can be represented solely by its linear component; the mean linear coefficients varied from .30 to .91 between the two tasks, while the analogous nonlinear averages were but .04 and .05.

Lower values for the observed and subjective complexity variables suggest more complex judgmental policies. While the average judge felt he was using more cues in the sociability task, he actually appeared to be using a more complex policy when predicting leadership. In any case, the average value for "insight" (the correlation between estimated and computed predictor weights) was considerably lower for predictions of leadership than for the other three

Table 2

Mean Values of the Judgmental Variables on the
First Replication of Each of the Four Tasks
(N=86)

Variable	Task			
	GPA	Leadership	Intelligence	Sociability
Linear Predictability	.92	.75	.94	.83
Consistency[a]	.87	.65	.90	.73
Linear Consistency[a]	.98	.93	.98	.95
Nonlinear Consistency[a]	.27	.28	.29	.24
Accuracy	.67	.13	-	-
Linear Accuracy	.91	.30	-	-
Nonlinear Accuracy	.05	.04	-	-
Observed Complexity	15.85	14.30	18.54	16.36
Subjective Complexity[a]	8.35	8.27	9.07	6.57
Insight[a]	.70	.36	.79	.69
Differentiation	1.92	1.84	2.07	1.60
Confidence	.61	.55	.58	.53

[a] These values, from the first occasion of each task, involve both replications.

tasks. Insight was highest for the same two tasks (intelligence and GPA) for which confidence was greatest and for which the most differentiation occurred.

Generality Analyses

The intercorrelations among the values for each judgmental variable, across replications within occasions, across occasions within tasks, and across tasks are available in Ramanaiah (1971).[4] For purposes of illustration, a few representative correlation matrices are included here as Tables 3, 4, and 5, and a summary of all of

[4]A quantitative index of generality is provided by the proportion of variance accounted for by the first unrotated principal factor, when these correlation matrices are factor analyzed. Findings based on such analyses are available in Ramanaiah (1971).

Table 3

Generality Correlations for Accuracy and for Linear Accuracy

(N = 86)

			Task						
			GPA				Leadership			
			Occasion I		Occasion II		Occasion I		Occasion II	
			Rep. A	Rep. B	Rep. A	Rep. B	Rep. A	Rep. B	Rep. A	Rep. B
GPA	I	A	-	.79**	.44**	.44**	.10	.15	.13	.10
		B	.81**	-	.40**	.31**	.04	.05	.02	.02
	II	A	.58**	.50**	-	.88**	-.02	-.08	-.18	-.01
		B	.56**	.44**	.93**	-	-.05	-.03	.11	.03
Leadership	I	A	.06	.03	.14	.15	-	.59**	.28**	.31**
		B	.08	.06	.01	.07	.52**	-	.46**	.42**
	II	A	.08	.01	-.11	-.14	.35**	.41**	-	.69**
		B	.03	.01	-.17	-.19	.29**	.53**	.80**	-

Note: Correlations among the accuracy scores are presented above the main diagonal, and
correlations among the linear components of accuracy are presented below the
diagonal.

**p ≤ .01

them is provided in Table 6. Table 3 provides the generality coefficients for judgmental accuracy (above the diagonal) and its linear component (below the diagonal). In this and subsequent tables, the two replications within each occasion are indexed as A and B, while the two occasions are indexed as I and II. Note that all of the across-replication within-occasion correlations, as well as all of the across-occasion within-task correlations, were quite high, while none of the across-task indices of generality were even statistically significant. Judgmental accuracy, as was hypothesized, shows considerable task specificity.

In contrast, the analogous coefficients for consistency and for its linear component, which are presented in Table 4, show substantial generality both within and across the tasks. The generality correlations for the linear component of consistency may well be attenuated due to restriction in range, since virtually every judge had very high, and thus very similar, linear consistency values.

Table 5 presents the generality coefficients for linear predictability (R_s) and for confidence. As one would expect, the values for linear predictability (Table 5) are quite similar to those for consistency (Table 4); in both cases, within-task

Table 4

Generality Correlations for Consistency and for Linear Consistency

(N=86)

		GPA I	GPA II	Leadership I	Leadership II	IQ	Sociability
GPA	I	–	.45**	.38**	.36**	.39**	.19
	II	.28**	–	.48**	.56**	.59**	.45**
Leadership	I	.36**	.28**	–	.67**	.40**	.46**
	II	.26*	.31**	.54**	–	.34**	.34**
Intelligence		.27**	.33**	.17	.26*	–	.44*
Sociability		.15	.29**	.21*	.06	.37**	–

Note: Correlations among the consistency scores are presented above the main diagonal, and correlations among the linear components of consistency are presented below the diagonal.

*p ≤ .05

**p ≤ .01

Table 7

Generality Correlations for Linear Predictability and for Confidence

(N=86)

		GPA				Leadership				Intelligence		Sociability	
		IA	IB	IIA	IIB	IA	IB	IIA	IIB	A	B	A	B
GPA	IA	-	.92	.34	.31	.44	.36	.26	.27	.38	.31	.14	.17
	IB	.99	-	.45	.40	.39	.33	.28	.29	.45	.39	.23	.26
	IIA	.81	.84	-	.86	.19	.26	.56	.54	.41	.48	.35	.37
	IIB	.82	.84	.99	-	.24	.32	.54	.54	.48	.50	.46	.42
Leadership	IA	.90	.88	.72	.73	-	.88	.55	.58	.24	.16	.31	.36
	IB	.89	.89	.74	.74	.99	-	.62	.64	.22	.18	.31	.34
	IIA	.73	.75	.90	.90	.79	.80	-	.95	.23	.25	.19	.29
	IIB	.69	.71	.85	.85	.73	.75	.91	-	.23	.25	.17	.26
Intelligence	A	.82	.84	.86	.85	.74	.75	.77	.72	-	.87	.50	.49
	B	.82	.84	.87	.86	.75	.77	.79	.73	.98	-	.50	.50
Sociability	A	.74	.76	.87	.87	.66	.68	.77	.73	.85	.86	-	.87
	B	.72	.75	.86	.86	.64	.65	.76	.71	.83	.84	.97	-

Note: Correlations among the linear predictability scores are presented above the main diagonal, and correlations among the confidence scores are presented below the diagonal. Correlations above 0.21 and 0.27 are statistically significant at probability levels of 0.05 and 0.01, respectively.

within-occasion correlations were very high, within-task across-occasion correlations were substantial, and most of the across-task correlations were statistically significant. The analogous values for judgmental confidence were extraordinary; of the 66 generality coefficients for confidence presented in Table 5, the lowest is .64. Judges clearly differ substantially in their confidence ratings, and this individual difference shows enormous generality across tasks of this kind.

Table 6, which summarizes the findings from these generality analyses, provides a direct comparison among the variables as potential judgmental styles. Interestingly, the variables that are computationally most simple fared more successfully in this generality contest than did the more complex ones. For example, the index

Table 6

The Generality of the Judgmental Variables:

A Summary of the Findings

Variable	Internal Consistency (Within-occasions, Within-tasks)		Temporal Stability (Across-occasions, Within-tasks)		Intertask Congruence (Across-occasions, Across-tasks)	
	Range	Mdn	Range	Mdn	Range	Mdn
*Linear Predictability	.86 - .95	.87	.31 - .64	.50	.14 - .56	.32
*Consistency	-	-	.47 - .67	.57	.19 - .59	.40
Linear Consistency	-	-	.28 - .54	.42	.06 - .37	.27
Nonlinear Consistency	-	-	.02 - .30	.16	.08 - .48	.18
Accuracy	.59 - .88	.74	.28 - .46	.41	-.18 - .15	.03
Linear Accuracy	.53 - .93	.80	.29 - .58	.47	-.19 - .15	.03
Nonlinear Accuracy	.01 - .47	.24	.00 - .39	.16	-.20 - .23	.05
Observed Complexity	.50 - .90	.77	.34 - .55	.41	-.04 - .37	.15
*Subjective Complexity	-	-	.52 - .63	.58	.38 - .72	.55
Insight	.82 - .92	.85	.28 - .40	.36	.05 - .53	.29
*Differentiation	.76 - .97	.94	.55 - .71	.66	.20 - .59	.36
*Confidence	.91 - .99	.99	.73 - .84	.80	.64 - .90	.77

* Variables selected as provisional judgmental styles.

based on a mean rating (confidence) showed extraordinarily high intertask congruence, followed by two of the three indices based upon rating dispersions (differentiation and subjective complexity) and the simple index of test-retest consistency. At the other extreme, the nonlinear indices, the index of observed complexity, and the accuracy variables all fared quite badly. On the basis of these findings it seems reasonable to conclude that two variables (confidence and subjective complexity) appear unusually promising as judgmental styles, while three more (differentiation, consistency, and linear predictability) merit study on a more provisional basis.

The variable "insight" might have been included in the latter set, since its generality appears equivocal on the basis of the findings summarized in Table 6. This quite complex variable has been operationalized as the correlation between: *(a)* the set of subjective weights when the judge is asked to formalize his prediction strategy by distributing 100 points among the various predictors, and *(b)* the set of computed relative weights based on the regression model of his judgments. As a result, the value of this index is affected both by the variance within the set of subjective weights and by the variance within the set of computed weights. The former, here termed "subjective complexity," has been shown to have considerable generality, and thus to be a provisional judgmental style in its own right. The latter, here termed "observed complexity," has been shown to have very limited generality across tasks. The minimal generality displayed by "insight" might well be an epiphenomenon, stemming solely from its sizeable artifactual relationship with subjective complexity (r = .60; p < .001) on the one hand, and perhaps with consistency (r = .53; p < .001) and/or linear predictability (r = .56; p < .001) on the other.

Correlations among the Provisional Styles

Consequently, all subsequent analyses will be limited to five provisional judgmental styles: linear predictability, consistency, subjective complexity, differentiation, and confidence. The intercorrelations among the composite indices

for each of these five variables are presented in Table 7. While the results of the previously-reported generality analyses were remarkably similar in the male and the female subsamples, some between-sample differences appeared for the first time at this stage of the investigation. Consequently, Table 7 provides the correlations separately within each of the subsamples, as well as for the total sample. As one would expect, in both subsamples consistency and linear predictability were extremely highly related, while confidence and differentiation were virtually unrelated to each other and to the other variables. On the other hand, subjective complexity was unrelated to any of the judgmental variables in the female subsample, while it was correlated about .60 with consistency and predictability in the male subsample.

Personality Scale Correlates of the Provisional Styles

This type of inter-sample difference increased dramatically when the provisional styles were correlated with the personality scale scores. Indeed, the pattern of personality scale correlates was radically different for the two subsamples. Moreover, monte carlo analyses within this data set indicated that the number of significant personality scale correlations that were obtained far exceeded the number to be expected by chance in the male subsample, while the number obtained in the female subsample was within the random range. Consequently, only the results based on the male subsample will be presented in this report.

For the male judges, those personality scales that were most highly correlated with the five provisional styles are listed in Table 8. Since linear predictability and consistency were so closely related, they produced a highly similar correlation pattern. These two variables were correlated quite substantially with the MMPI scales constructed by Welsh to mark the second *(R)* and the third *(C)* factors of that inventory, and to an analogous scale (called "Person Orientation" [PO] by Nichols and Schnell [1963]) constructed to measure the second factor of the

Table 7

Intercorrelations Among the Five Provisional Styles

	Linear Predictability		Consistency		Subjective Complexity		Differentiation		Confidence	
	M	F	M	F	M	F	M	F	M	F
Predictability	–		.89		.28**		.22*		-.04	
Consistency	.94**	.86**	–		.32**		.30**		.01	
Complexity	.58**	.06	.61**	.12	–		.22*		-.03	
Differentiation	.29	.16	.30	.26	.37*	.10	–		.12	
Confidence	-.06	-.03	-.10	.06	-.17	.06	.08	.14	–	

Note: Correlations listed above the main diagonal are based on the total sample for whom personality data were available (N=83); correlations listed below the diagonal are based on the male (M) subsample (N=35) and the female (F) subsample (N=48).

*p < .05

**p < .01

CPI. Judgmental consistency was positively related to the MMPI scales for Hypochondriasis (Hs), Depression (D), and Social Introversion (Si), as well as to college grade point average and the CPI Achievement via Independence (Ai) scale, which was constructed to predict college grades. Consistency was negatively related to the CPI scale constructed by Hase and Goldberg (1967) to measure Need for Affiliation (nAf), the Personnel Director scale from the SVIB, and the total amount of past experience in diverse activities as measured by a 320-item Reported Behavior Inventory (Goldberg & Jones, 1969).

Judgmental confidence was negatively related to indecisiveness (as indexed by the number of items omitted on the MMPI), the Flexibility (Fx) scale of the CPI, and masculinity as measured by scales from the MMPI (Mf), the CPI (Fe) the EPPS (Het), and the SVIB (M-F). This feminine orientation was further reflected in three significant SVIB correlates, namely the occupational scales for Psychologist, Musician, and Minister. Judgmental confidence was positively related to a self-rating of personal popularity from the biographical inventory, and negatively related to college grades. Since judgmental confidence and judgmental consistency were themselves unrelated (r = -.10; see Table 7) and since both were significantly associated with college GPA (r = .41 and -.46), it is possible to predict scholastic performance by these two variables in this small sample with a multiple correlation of nearly .60, a precision that is usually only possible with a measure of past performance (high school GPA) and a scholastic aptitude test. Clearly this finding should be replicated.

Table 8

The Highest Correlates of the Provisional Styles:

Male Subsample (N=35)

Predictability (a) and Consistency (b)			Complexity	Differentiation	Confidence
	a	b			
MMPI: R	(.47)	(.52)	EPPS: Dom (-.43)	CPI: nAc (-.52)	MMPI: "?" (-.53)
MMPI: C	(.45)	(.48)	EPPS: Suc (.38)	SVIB: Senior CPA(-.51)	MMPI: Mf (-.49)
CPI: Po	(-.48)	(-.48)	SVIB: R.E. Sales. (.37)	SVIB: Acct. (-.44)	Popularity (.48)
Experience (RBI)	(-.40)	(-.47)	SVIB: Chemist (-.37)	SVIB: Int. Mat. (-.41)	College GPA (-.46)
Freshman GPA	(.40)	(.45)		SVIB: Off. Mgr. (-.41)	CPI: Fe (-.42)
MMPI: Hs	(.42)	(.44)		SVIB: Aut.-Jrn. (.37)	SVIB: Psychol. (-.40)
MMPI: D	(.45)	(.44)		SVIB: Pers. Dir.(-.37)	SVIB: Musician (-.39)
MMPI: Si	(.40)	(.43)		CPI: Soc (-.36)	SVIB: Minister (-.36)
CPI: nAf	(-.41)	(-.43)		CPI: Sp (-.35)	EPPS: Het (-.36)
SVIB: Pers. Dir.	(-.36)	(-.43)		CPI: Sa (-.35)	SVIB: M-F (.34)
College GPA	(.35)	(.41)		CPI: Cm (-.35)	MMPI: Mp (.34)
CPI: Ai	(.39)	(.41)		MMPI: R (.35)	CPI: Fx (-.33)

Judgmental differentiation was associated positively with the SVIB occupational scale for Author-Journalist, and negatively with a host of SVIB and CPI scales. Specifically, "cautiousness" or lack of extremeness in judgment was related to such SVIB occupational scales as Senior Certified Public Accountant, Accountant, Office Manager, and Personnel Director, as well as the SVIB scale for Interest Maturity, which reflects interests more like older than younger workers. Differentiation was also negatively related to two CPI scales developed by Hase and Goldberg (1967) to measure Need for Achievement (nAc) and Sociability (Soc) and to a number of the standard CPI scales, including Social Presence (Sp), Self-Acceptance (Sa), and Communality (Cm).

The significant correlates of *subjective complexity* were neither as numerous nor as coherent as those for the other four provisional styles. Judgmental complexity was positively related to the EPPS scale for Succorance (Suc) and the SVIB scale for Real Estate Salesman, and negatively related to the EPPS scale for Dominance (Dom) and the SVIB Chemist scale. For a more complete accounting of the correlates of these provisional styles, see Ramanaiah (1971).

Discussion

The findings from the present study should provide an empirical bridge between some important styles of human decision-making and other aspects of personality structure. The reader's assessment of the strength of that bridge will depend on his willingness to generalize beyond this sample of judges and this set of judgment tasks. For unfortunately, scientists who have previously studied individual differences in judgmental processes heretofore have not sought to ascertain the trans-situational generality of those individual differences they have explored.

Indeed, with but a few exceptions (e.g., Block & Petersen, 1955), such individual differences have been investigated solely from the vantage point of two quite restricted methodologies. In one, similarity ratings among pairs of stimuli (e.g., a small set of trait terms, names of a few political leaders or nations) have been analyzed via one of the multidimensional scaling algorithms that permit individual differences to emerge; initially most investigators (e.g., Messick & Kogan, 1966; Pedersen, 1965; Walters & Jackson, 1966; Wiggins & Hoffman, 1968) used the technique pioneered by Tucker and Messick (1963), while more recent efforts (e.g., Jones & Young, 1972; Sherman, 1972; Sherman & Ross, 1972; Wish, Deutsch, & Biener, 1970) have employed a similar procedure developed by Carroll and Chang (1970).

Alternatively, using more natural types of judgments (e.g., preferences, trait attributions, clinical diagnoses), the relationships among judges have been analyzed via an obverse factor analysis of the inter-judge cross-product, covariance, or correlation matrix; sets of judges clustering together on the resulting "subject factors" are then compared on personality or other variables (e.g., Posavac, 1971; Shikiar, Fishbein, & Wiggins, 1974; Wiggins, 1971; Wiggins, Hoffman, & Taber, 1969). Recently, both types of studies have been reviewed by Wiggins (1973), who concluded with a plea for future research on the "basic parameters of individual differences in human judgment: the reliability, internal consistency, and transituational generality of such individual differences" (p. 137). This was precisely the focus of the present investigation.

When a complex problem is framed in starkly polar alternatives (e.g., heredity *versus* environment, content *versus* style), scientific illumination is unlikely. In terms of today's most popular polarity, the trait versus situation controversy (Mischel, 1968), the present findings should displease extremists on the right and the left. For clearly some individual differences (e.g., confidence) appear to possess substantial generality, while others (e.g., accuracy) do not. If one genuinely expects all individual differences to generalize across situations, or none to do so, one will be disappointed. (As test-wise students can affirm, statements which include such terms as

"all," "always," "none," or "never" are rarely true.)

In the present case, of the total set of individual differences originally considered, roughly half appear to possess sufficient generality to warrant further investigation as possible judgmental styles. Specifically, the indices of linear predictability, consistency, subjective complexity, differentiation, and confidence showed promise, while the nonlinear indices, the accuracy measures, and the indices of observed complexity and "insight" did not. Retrospectively, the differences between these two classes of variables seem sensible: By and large, measures reflecting the central tendency or dispersion of a judge's responses (e.g., differentiation), or of his evaluation of them (e.g., confidence), appear to possess more generality than measures that are less direct (e.g., observed complexity) or that reflect an inherent interaction between characteristics of the task and those of the judge (e.g., accuracy).

The six provisional styles are remarkably independent, with but two exceptions: a) Linear predictability (R_s) and consistency are necessarily related. While a judge might conceivably employ a non-linear policy (e.g., a U-shaped relation between a cue and a judgment) in a completely consistent fashion (thus generating low linear predictability and high consistency), the reverse (high linear predictability and low consistency) is not possible. Moreover, since a host of previous studies have shown that for most judges and most tasks a linear model captures virtually all of the reliable judgmental variance (Goldberg, 1968; Slovic & Lichtenstein, 1973), consistency merely serves to limit linear predictability and the correlation among these two indices must approach unity. b) In the present study, predictability and consistency were related to subjective complexity in the male, though not in the female, subsample. Scatterplots of these two relationships show that the effect of sex as a moderator variable was not an artifact produced by one or two deviant subjects in either subsample. On the other hand, one must not make too much of this correlational differ- ence prior to replication in other samples.

Some readers may have expected to find a substantial relationship between confidence and differentiation, on the grounds that more extreme judgments should reflect more certainty than less extreme ones. Indeed, for most judges intra-individual confidence or "individual certainty" (Little, 1961) is highly related to rating extremeness. However, individual differences in judges' mean confidence ratings, their "generalized confidence" (Little, 1961) are typically related neither to accuracy (Goldberg, 1959) nor to their mean level of differentiation. Little (1961) has shown the importance of distinguishing between inter-individual differences in mean confidence and intra-individual differences in certainty (i.e., deviations from the individual's mean). In his classic study of both types of confidence, Little found substantial reliability for individual differences in mean confidence, reflected in test-retest correlations around .70 over a four to six week interval, and equally substantial generality ($r = .72$) across two different judgmental tasks. In addition, Little found no relationship between individual differences in mean confidence and consistency. All of these findings have been replicated in the present study.

In one of the few studies of the correlates of individual differences in mean confidence, Block and Petersen (1955) related confidence on a psychophysical task to staff evaluations made during a living-in assessment program at Berkeley's Institute of Personality Assessment and Research. In a sample of 53 military officers, high judgmental confidence was associated with Q-sort ratings of rigidity and dogmatism, while low confidence was associated with ratings of introspection and self-abasement. In the present study, judgmental confidence was highly related to self-reported popularity and to diverse personality scale manifestations of "masculine" (as opposed to "feminine") interests; lack of confidence was here associated with indecisiveness in answering MMPI items and "flexibility" as measured by the CPI. Unfortunately, since correlations with personality scale scores were not reported by Block and Petersen (1955), it is

not possible to examine the extent of their convergence between the two studies.

The pattern of scales associated with individual differences in judgmental consistency and linear predictability (basically those loading highly on the second major MMPI factor [R] and the first major CPI factor [PO]) are virtually identical to those related to test-retest response consistency on three personality inventories in the larger group of subjects from which this sample was recruited (Goldberg & Jones, 1969). That is, response consistency as measured by the proportion of personality inventory items changed upon retest after a one-month interval was quite highly related to judgmental consistency and linear predictability in radically different types of tasks administered almost a year later. And, as a consequence, the general pattern of consistency correlates was quite similar for both item responses and judgmental tasks. Clearly this intriguing finding now demands replication in other populations.

Just as psychometric response consistency seems to be related to policy consistency on diverse judgmental tasks, so judgmental differentiation should be related to the one uncontroversial psychometric response style, namely extremeness response set (Hamilton, 1968). In the present study, judgmental "cautiousness" was highly correlated with the type of compulsivity indexed by various SVIB accountancy scales. An important task for future investigators must be to relate the two different manifestations of differentiation in the same sample of subjects and seek to replicate the personality correlates here discovered.

References

Anderson, N. H. A simple model for information integration. In R. P. Abelson, et al. (Eds.), *Theories of cognitive consistency: A Sourcebook.* Chicago: Rand McNally, 1968. Pp. 731-743.

Block, J., & Peterson, P. Some personality correlates of confidence, caution, and speed in a decision situation. *Journal of Abnormal and Social Psychology,* 1955, 51, 34-41.

Campbell, D. T., & Fiske, D. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin,* 1959, 56, 81-105.

Carroll, J. D., & Chang, J. J. Analysis of individual differences in multidimensional scaling via an *N*-way generalization of "Eckart-Young" decomposition. *Psychometrika,* 1970, 35, 283-319.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. *The dependability of behavioral measurements: Theory of generalizability for scores and profiles.* New York: Wiley, 1972.

Crow, W. J., & Hammand, K. R. The generality of accuracy and response sets in interpersonal perception. *Journal of Abnormal and Social Psychology,* 1957, 54, 384-390.

Edwards, W. Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing. *Journal of Mathematical Psychology,* 1965, 2, 312-329.

Goldberg, L. R. The effectiveness of clinicians' judgments: the diagnosis of organic brain damage from the Bender-Gestalt test. *Journal of Consulting Psychology,* 1959, 23, 25-33.

Goldberg, L. R. Simple models or simple processes? Some research on clinical judgments. *American Psychologist,* 1968, 23, 483-495.

Goldberg, L. R. Man versus model of man: A rationale, plus some evidence, for a method of improving on clinical inferences. *Psychological Bulletin,* 1970, 73, 422-432.

Goldberg, L. R., & Jones, R. R. The reliability of reliability: The generality and correlates of intra-individual consistency in responses to structured personality inventories. *Oregon Research Institute Research Monograph,* 1969, Vol. 9, No. 9.

Hamilton, D. L. Personality attributes associated with extreme response style. *Psychological Bulletin,* 1968, 69, 192-203.

Hammond, K. R., Hursch, C. J., & Todd, F. J. Analyzing the components of clinical inference. *Psychological Review,* 1964, 71, 438-456.

Hase, H. D., & Goldberg, L. R. Comparative validity of different strategies of constructing personality inventory scales. *Psychological Bulletin,* 1967, 67, 231-248.

Hoffman, P. J. The paramorphic representation of clinical judgment. *Psychological Bulletin,* 1960, 57, 116-131.

Johnson, L. C. Generality of speed and confidence in judgment. *Journal of Abnormal and Social Psychology,* 1957, 54, 264-266.

Jones, L. E., & Young, F. W. Structure of social environment: Longitudinal individual differences scaling of an intact group. *Journal of Personality and Social Psychology,* 1972, 24, 108-121.

Kahneman, D., & Tversky, A. On the psychology of prediction. *Psychological Review,* 1973, 80, 237-251.

Little, K. B. Confidence and reliability. *Educational and Psychological Measurement,* 1961, 21, 95-101.

Messick, S., & Kogan, N. Personality consistencies in judgment: Dimensions of role constructs. *Multivariate Behavioral Research,* 1966, 1, 165-175.

Mischel, W. *Personality and assessment.* New York: Wiley, 1968.

Naylor, J. C., & Wherry, R. J., Sr. The use of simulated stimuli and the "Jan" technique to capture and cluster the policies of raters. *Educational and Psychological Measurement,* 1965, 25, 969-986.

Nichols, R. C., & Schnell, R. R. Factor scales for the California Psychological Inventory. *Journal of Consulting Psychology,* 1963, 27, 228-235.

Pedersen, D. M. The measurement of individual differences in perceived personality-trait relationships and their relation to certain determinants. *Journal of Social Psychology,* 1965, 65, 233-258.

Posavac, E. J. Dimensions of trait preferences and personality type. *Journal of Personality and Social Psychology,* 1971, 19, 274-281.

Ramanaiah, N. V. A search for stylistic components of judgments. Unpublished doctoral dissertation, University of Oregon, 1971.

Sherman, R. C. Individual differences in perceived trait relationships as a function of dimensional salience. *Multivariate Behavioral Research,* 1972, 7, 109-129.

Sherman, R. C., & Ross, L. B. Liberalism-conservatism and dimensional salience in the perception of political figures. *Journal of Personality and Social Psychology,* 1972, 23, 120-127.

Shikiar, R., Fishbein, M., & Wiggins, N. Individual differences in semantic space: A replication and extension. *Multivariate Behavioral Research,* 1974, 9, 201-209.

Shrauger, S., & Altrocchi, J. The personality of the perceiver as a factor in person perception. *Psychological Bulletin,* 1964, 62, 289-308.

Slovic, P., & Lichtenstein, S. Comparison of Bayesian and regression approaches to the study of information processing in judgment. In L. Rappoport & D. A. Summers (Eds.), *Human judgment and social interaction.* New York: Holt, Rinehart & Winston, 1973, Pp. 16-108.

Tucker, L. R. A suggested alternative formulation in the developments by Hursch, Hammond, and Hursch, and by Hammond, Hursch, and Todd. *Psychological Review,* 1964, 71, 528-530.

Tucker, L. R., & Messick, S. An individual differences model for multidimensional scaling. *Psychometrika,* 1963, 28, 333-367.

Tversky, A., & Kahneman, D. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology,* 1973, 5, 207-232.

Walters, H. A., & Jackson, D. N. Group and individual regularities in trait inference: A multidimensional scaling analysis. *Multivariate Behavioral Research,* 1966, 1, 145-163.

Wiggins, N. Individual differences in diagnostic judgments of psychosis and neurosis from the MMPI. *Educational and Psychological Measurement,* 1971, 31, 199-214.

Wiggins, N. Individual differences in human judgment: A multivariate approach. In L. Rappoport & D. A. Summers (Eds.), *Human judgment and social interaction.* New York: Holt, Rinehart & Winston, 1973. Pp. 110-142.

Wiggins, N., & Hoffman, P. J. Dimensions of profile judgments as a function of instructions, cue-consistency, and individual differences. *Multivariate Behavioral Research,* 1968, 3, 3-20.

Wiggins, N., Hoffman, P. J., & Taber, T. Types of judges and cue-utilization in judgments of intelligence. *Journal of Personality and Social Psychology,* 1969, 12, 52-59.

Wish, M., Deutsch, M., & Biener, L. Differences in conceptual structures of nations: An exploratory study. *Journal of Personality and Social Psychology,* 1970, 16, 361-373.

Author's Address

Lewis R. Goldberg, Oregon Research Institute, P.O. Box 3196, Eugene, Oregon 97403.