

# A signal detection analysis of the rod and frame test<sup>1</sup>

JOHN R. SCHUCK,<sup>2</sup> BOWLING GREEN STATE UNIVERSITY, HENRY A. CROSS, COLORADO STATE UNIVERSITY, AND DAVID H. MILLS, UNIVERSITY OF MARYLAND

*To help determine just what Witkin's Rod and Frame Test measures, the task was modified by presenting the luminous rod in one of two fixed positions, either vertical or not vertical, within a tilted frame. S responded on a four-category scale representing his certainty that the rod was vertical. Receiver operating characteristic (ROC) curves were calculated for each of 17 Ss and compared to traditional measures of their field dependence-independence. The results showed that Witkin's test is not contaminated to any great extent by response bias. However, the ROC analysis did reveal significant differences between field dependents and independents, not only in their sensitivity to verticality, but also in their characteristic manner of responding.*

Several decades ago, Witkin and his colleagues began a research program dealing with the way people orient themselves in space. One of their initial tests was the Rod and Frame. In this task, S sits in a darkened room, facing a luminous rod within a square, luminous frame. Initially, both rod and frame are presented in a tilted position. Gradually, E moves the rod until S reports it as vertical. Some Ss report that the rod is vertical when its orientation is close to that of the tilted frame. These Witkin calls field dependent. Others set the rod close to the true vertical and are referred to as field independent.

There is some question, however, about the extent to which performance on the Rod and Frame task is contaminated by nonperceptual factors. For example, Gross (1959) led one group of Ss to believe that a special lens was being used to increase the ambiguity of the position of the rod; in reality clear glass was used. Afterward, all Ss were given a checklist and told to double-check the term that best described their feelings during the experiment. This set for uncertainty resulted in a significant increase in field dependency. Furthermore, field-dependent Ss checked "uncertainty" significantly more often than did the field-independent Ss. Linton (1965) studied the relations among several measures of field dependence and the social conformity of her Ss. One measure of conformity was obtained by having S write down his report of autokinetic

movement in a darkened room. With him in the room was another person who pretended to be a fellow S, but who actually was E's confederate. On some trials, the confederate reported aloud, and S's conformity to the confederate's report was recorded. Results showed that field dependence was significantly correlated with this kind of conformity.

These findings are cited by Witkin, Dyk, Faterson, Goodenough, and Karp (1962) in support of their view that perceptual ability is but one aspect of a much broader dimension that also involves consistent cognitive, intellectual, social, and affective styles of behavior. In fact, the Witkin group presents additional evidence to support the notion that field dependents, relative to independents, rely more on people (especially authority figures) for guidance and support and are more attentive to the faces of those around them.

Obviously, an alternative interpretation is feasible. Perhaps performance on the Rod and Frame Test is not primarily the result of a perceptual process but is partially or even largely the artifactual consequence of such nonperceptual factors as habitual styles of dealing with other people, particularly in ambiguous situations. In order to explicate this notion, let us consider the problem from the vantage point of signal detection theory.

Assume, for the moment, that the Rod and Frame Test can be modified so that "hit" and "false alarm" rates may be calculated. A hit would be S's report that the rod was vertical when it was vertical. A false alarm would be S's report that the rod was vertical when it actually was tilted. Assume, for the sake of argument, that the field-dependent S is just as sensitive to verticality as the independent S. His dependence on other people and their opinions, however, particularly in an ambiguous experimental situation, leads him to "hedge his bets" in reporting the position of the rod. He does not want to deviate too much from E's expectations, but he is not too sure just what E does expect; it may be unclear to S, for instance, just what the luminous frame has to do with the experiment. Hence, the field-dependent S might be prone to deliberately misreport a certain proportion of trials in order not to be too extreme in

weighting or in failing to weight the importance of the frame. In this case, areas under the ROC functions of field-dependent Ss should be artifactually reduced, but the ROC functions themselves should still either overlap or lie above the chance diagonal. If, however, the field dependent actually perceives the tilted rod as vertical, his ROC function should lie below the chance diagonal.

The primary purpose of this study was to apply a signal-detection analysis to the Rod and Frame task. One of our goals was to measure sensitivity to the vertical independently of any nonperceptual factors influencing the setting of the rod. Another was to quantify possible differences between field dependents and field independents with respect to their response strategies in this situation. Moreover, we wanted to compare our new measure of sensitivity to the vertical with four traditional measures of field dependence.

## METHOD AND PROCEDURE

### Subjects

The Ss were 13 girls and 4 boys, ranging in age from 8 to 11 years. All were volunteers who came from one of several middle-class neighborhoods of Ames, Iowa, and who had heard about the experiment by word of mouth. Each S was paid 50 cents a session. Participation apparently was partly for the money and partly for the prestige of being taken to the University for psychological experimentation. Of the first 12 Ss to serve in the study, only 4 were field independent, as measured by the initial test with the standard Witkin apparatus. In order to insure that about half of the Ss in the study were independent, an additional 12 Ss were given the standard Witkin test, and the 5 members of this group who showed independence were selected for further study. According to this pretest, 7 girls and 2 boys were independents, with an average age of 9.8 years. Six girls and 2 boys were dependents, with an average age of 9.3 years. The mean absolute error of the rod setting for the group of independents was 3.66 deg, with a standard deviation of 1.46. For the dependents, the mean absolute error was 16.59 deg, with a standard deviation of 4.73.

## Apparatus

A standard Witkin rod and frame was used for testing during the Witkin pretest and posttest. The frame was 40 x 40 x 1 in. and was permanently tilted 28 deg to S's left. The rod was 36 x 1 in. The apparatus was painted flat black and the surfaces of both the frame and the rod were coated with luminous paint. S was seated behind a screen 8 ft from the apparatus. A portion of the screen could be lowered by E, allowing S to view the apparatus through a round aperture, 6 in. in diam. The screen was also painted flat black.

A specially designed rod and frame, with dimensions similar to those of the standard apparatus, was used to collect the data by the signal-detection method. The frame was constructed by mounting strips of electroluminescent Tape-lite (Sylvania Co.) on a large circular panel that permitted quick adjustment of position and angle of tilt. The rod was also of Tape-lite. A Hunter timer was used to control the duration of exposure, and the input voltage was adjusted so that the luminance of the Tape-lite was subjectively equivalent to the luminous paint on the standard apparatus. As before, a screen prevented S from observing the adjustments of the rod and frame during the intertrial interval.

## Procedure

During pretesting and posttesting, S was led, with eyes closed, into a totally darkened room and seated behind the screen. He was given the following instructions: "Do you have a flagpole at your school? Good; then you know what straight up and down is. In just a little while, I am going to show you a rod that is inside a frame. The rod will be tipped over, but I will move it in short steps and I want you to tell me when I have the rod so that it is exactly like the flagpole at your school. Each time I move the rod, say either 'yes' or 'no.' Say 'no' if the rod is still tipped over but 'yes' if it is straight up and down." After these instructions, E tilted the rod 28 deg to the S's right, lowered the screen in front of S, and began

adjusting the rod to the vertical in 3-deg steps. Adjustment continued until S responded with "yes." On subsequent trials, the initial tilt of the rod varied between 10 and 28 deg, except that the rod was never set inside any of S's estimations. The pretest consisted of 24 trials, and the posttest was identical except that it involved only 16 trials. The experimental room was dark throughout the period of exposure and the intertrial interval, which was approximately 45 sec.

The task following the pretest was modified so that receiver operating characteristic (ROC) curves could be obtained for each S by the rating method described in Green and Swets (1966). Basically, the procedure involved presenting the rod in one fixed position, either vertical or not vertical, for a short observation interval and having S respond on a four-category scale (yes-sure, yes-not sure, no-not sure, no-sure), representing his certainty that the rod was vertical. A *hit* was S's indication of verticality when the rod was actually vertical. A *false alarm* was when S indicated verticality but the rod was not actually vertical.

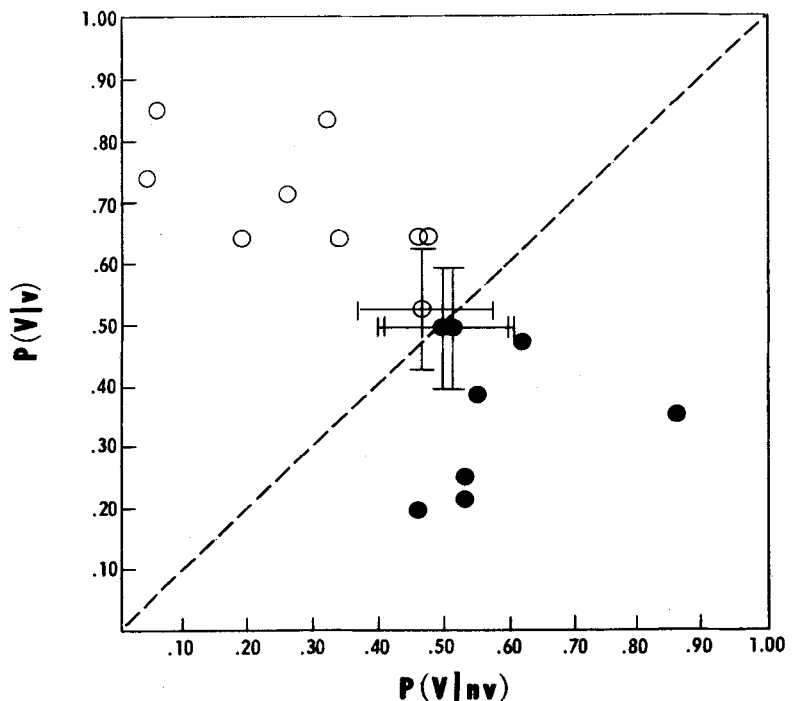
Testing during this period involved five daily sessions of 48 trials each. On each day, Ss received three blocks of 16 trials each. Each block of trials was separated by a 10-min rest interval. Instructions were as follows: "Now, your task is going to be very much like it was before. That is, you are to tell me when the rod is straight up and down like the flagpole at your school. This time, however, you will see the rod

for only a brief time. Immediately after the rod disappears you are to say either 'yes' or 'no.' Say 'yes' if the rod was straight up and down and 'no' if it was slightly tipped over. Half of the time, i.e., on half of the trials, the rod will be straight up and down and half of the time, it will be slightly tipped. In addition, I want you to tell me how sure you were of your statement of 'yes' and 'no.' After you say 'yes' or 'no,' please say either 'sure' or 'not sure.' Say 'sure' when you feel quite confident of your answer and say 'not sure' when you don't feel confident."

A pilot study, involving four Ss, was used to determine appropriate values for the tilt of the rod, tilt of the frame, and trial duration. The rod and frame were lighted for 2 sec in an otherwise darkened room. On any one trial, the frame was tilted 8 or 28 deg, either to the left or to the right of vertical. Whenever the rod was tilted, it was always set 6 deg from the vertical, in the same direction as the tilt of the frame. Thus, there were eight possible combinations of rod positions (vertical or not vertical), frame setting (8 or 28 deg), and direction of tilt (right or left). These eight possible combinations were repeated twice in a randomized sequence to make up each block of 16 trials.

The S was allowed to inspect the apparatus, with the frame not tilted, before the initial trials of Day 1 and was shown the exact rod settings that would be used. He was not shown the frame settings, but he was told that the frame would be adjusted, i.e., tilted, from time to time

Fig. 1.  $P(V|v)$  represents the proportion of "vertical" responses when the rod was actually vertical and  $P(V|nv)$  the proportion of "vertical" responses when the rod was actually not vertical. The nine Ss classified as field independent on the Witkin pretest are indicated by open circles. Closed circles indicate those classified as field dependents on the pretest. Bands for the 95% confidence intervals are shown for the three data points that did not differ significantly from chance performance.



during the test. During this period, E again emphasized that on half of the trials the rod would be vertical and that on the other half of the trials it would be tipped. It was further emphasized that since this was not an easy task, S should use each category of "sure" and "not sure" about half of the time.

During the first session of 48 trials, an attempt was made to shape the verbal responses of S in harmony with the instructions. After each block of 16 trials, S was told that he was going well but that he needed to say "yes" or "no" (as the particular case indicated) more often. He was also told this with reference to the use of "sure" and "not sure." Such instructions were given only if S had used a particular designation less than six times during the 16 trials. If S was already performing to criterion in this respect, he was simply told that he was doing well and that he should continue to respond in this fashion. At the beginning of Day 2, S was told that he would no longer be given information about how well he was performing but that he was to keep the instructions in mind at all times. At this point, S was asked to verbalize the instructions for E and was then presented with his first block of trials. This procedure, wherein S verbalized the instructions, was repeated each experimental session.

After the 5 days of this procedure, all Ss were given the posttest with Witkin's traditional method. Then the following tests were administered: the block-design and picture-completion subtests of the WISC, the embedded figures test, and the sophistication-of-body-image test; all of these are measures that Witkin and his colleagues have found to be highly correlated with their version of the rod and frame test (Witkin, Dyk, Faterson, Goodenough, & Karp, 1962).

## RESULTS

In order to test statistically for sensitivity to the vertical, one data point in the unit square was calculated for each S, based upon his yes-no responses and ignoring his confidence ratings. Then the two 95% confidence bands (one for the proportion of hits, the other for the proportion of false alarms) were calculated for this data point. Significant deviation from chance performance was indicated whenever neither confidence band intersected the positive diagonal.

Using this criterion, eight of the nine Ss who were classified as field independent on the Witkin pretest obtained data points significantly above the chance diagonal. Six of the eight Ss initially classified as field dependent obtained data points

Table 1  
Intercorrelations of the ROC Measures with Other Tests

	Field Independence or Psychological Differentiation					
	(-) ROC Area	Witkin Pretest	Witkin Posttest	Embedded Figures	(-) WISC Subtests	Body Image
(-) ROC Area	1.00	.82	.73	.66	.71	.65
Witkin-Pretest		1.00	.89	.72	.81	.66
Witkin-Posttest			1.00	.63	.68	.57
Embedded Figures				1.00	.83	.84
(-) WISC Subtests					1.00	.84

significantly below the chance diagonal. The confidence ellipses for the remaining three Ss overlapped the diagonal. These data are shown in Fig. 1.

For the remainder of this paper, Ss will be classified according to the ROC analysis rather than on the basis of the Witkin pretest. Those with data points significantly above the positive diagonal of the unit square will be classified as *independents*, those with data points significantly below, as *dependents*, and the remaining, as *others*.

The data points shown in Fig. 1 were obtained by combining the data from both the 8- and 28-deg settings of the frame. Individual curves were then obtained by the standard cumulative procedure for calculating such functions from rating-scale data. Then the area under the ROC curve was calculated for each S as a measure of that S's sensitivity to verticality. For a brief description of the rating method, see Hake and Rodwan (1966).

Table 1 shows the intercorrelations of the ROC sensitivity measures and the scores from the other tests of field dependence. As a check on the reliability of the sophistication-of-body-image test, it was not only scored in the usual way (Witkin et al, 1962) but also rank-ordered by two persons not familiar with the Ss. The rank-order correlation between judges was high ( $\rho = .79, p < .01$ ).

Since the range of scores on the body-image test was quite restricted, rank-order correlations were also calculated for all comparisons involving this test. The only comparisons in which the rank correlation changed the size of the coefficient to any extent were those involving ROC area (.71 instead of .65) and the Witkin posttest (.68 instead of .57). All of the coefficients listed in Table 1 are significant ( $p < .05$ ).

In order to compare performances under the two different tilts of the frame, two additional ROC functions were calculated for each S. One curve for each S was based solely upon those trials in which the frame was tilted 8 deg. The other curve included only data from those trials in which the frame was tilted 28 deg. Otherwise, the ROC curves and sensitivity measures were

calculated in the same way as those for the combined data.

The mean sensitivity scores and 95% confidence intervals for the 8-deg condition were  $.71 \pm .14$ ,  $.11 \pm .08$ , and  $.28 \pm .04$  for the independents, dependents, and others, respectively. Those for the 28-deg condition were  $.78 \pm .13$ ,  $.48 \pm .07$ , and  $.72 \pm .09$ , respectively.

Figure 2 shows how the Ss distributed their responses, on the average, across the four-category rating scale and suggests the existence of systematic differences among the rating functions for independents, dependents, and others. In order to provide some statistical evidence concerning this question, the data were further analyzed by a method used previously by Shontz (1967).

This analysis dealt with the 2 by 4 response matrices obtained for each S, one for the 8-deg frame setting and the other for the 28-deg setting. In each matrix, the two columns represented the vertical and tilted positions of the rod, and the four rows represented the rating categories used by each S: yes-sure, yes-not sure, no-not sure, no-sure. For each matrix, one column was inverted so that the "yes-sure" cell for the "vertical" column was adjacent to the "no-sure" cell of the "tilted" column, etc., and the a chi square was calculated. In the analysis, low values of chi square indicated that the S's criteria and confidence judgments when the rod was vertical were similar to those he used when it was not. Anything causing the S to distribute his confidence judgments differently in the two rod conditions would elevate the value of chi square. Thus, each chi square reflected the extent of the discrepancy between the way S used the rating scale when the rod was vertical and the way he used it when the rod was not vertical. Two discrepancy scores were obtained for each S by taking the square root of his chi-square values. Mean discrepancy scores for the 8-deg frame setting were 4.70, 5.98, and 5.67 for independents, dependents, and others, respectively. The respective means for the 28-deg rod setting were 3.15, 8.02, and 5.30.

A 3 by 2 analysis of variance (3 dependence categories by 2 frame settings)

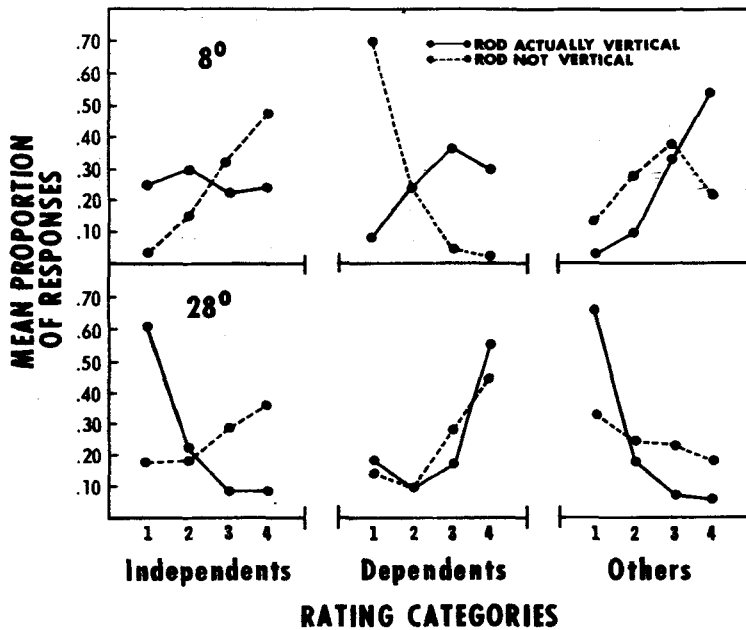


Fig. 2. Mean proportion of responses across the rating categories for the three classifications of Ss. Rating Category 1 represents yes-sure; 2, yes-not sure; 3, no-not sure; and 4, no-sure.

the dependents tended to differentiate successfully the vertical from the tilted position of the rod only when the frame was tilted 8 deg. This, of course, was the only condition in which the tilt of the rod (6 deg) approximated that of the frame, and so such a finding makes considerable sense from Witkin's point of view. A simple decision rule adequately accounts for the behavior of the dependents: when the rod and frame are about equally tilted, call the rod vertical; for all other conditions, call the rod "not vertical."

Figure 2 also suggests that the independents distributed their ratings in an almost rectangular distribution when the rod was vertical and the frame was tilted 8 deg. For the 28-deg tilt of the frame, however, their rating function reflects efficient and accurate performance for the vertical-rod condition. Conversely, when only the tilted rod is considered, the slopes of the rating functions for independents suggest a decrement in performance on the 28-deg as compared to the 8-deg frame condition. To summarize this aspect of the data in cognitive language, the independents appeared more confident of their judgments of the vertical rod when the frame was tilted 28 deg, but they were less confident of their judgments of the tilted rod in this condition.

Obviously, then, the perceptual judgments of the independents in this situation were not independent of the field, i.e., the frame. For both dependents and independents, the position of the frame had systematic effects upon performance. For dependents, this context seriously affected the overall veridicality of their judgments. For independents, it did not.

The three Ss classified as "others" were placed in this category because their ROC functions, based on the combined data, did not differ significantly from chance. Figure 2 indicates, however, that their performance was far from systematic. When the frame was tilted 8 deg, these Ss systematically rated the vertical rod as "not vertical." When the frame was tilted 28 deg, they were quite accurate in their rating of the vertical rod.

These kinds of inferences are supported by the mean sensitivity measures and confidence intervals calculated for the frame conditions separately. Only one of these confidence intervals included the hypothetical chance value of .50. This was for the dependents' performance when the

was applied to these scores with the results shown in Table 2. The F value for the dependency categories was 6.03 ( $p < .05$ ), and that for the interaction between frame setting and dependency approached significance ( $F = 3.31$ ,  $p < .10$ ). The mean discrepancy scores suggest that only the dependents and the independents contributed to the interaction, the independents having their highest discrepancy scores when the frame was set at 8 deg, the dependents obtaining their highest discrepancy scores with the 28-deg setting. In order to test this post hoc hypothesis, the Studentized Range Statistic was used to set up the 95% confidence interval about the difference representing this interaction, that is,  $(4.70 + 8.02) - (3.15 + 5.98)$ . The resulting confidence interval was  $3.59 \pm 3.28$ , thus indicating a significant interaction.

#### DISCUSSION

These results do not support the notion that field dependence is artifactual and explainable in terms of response bias. Disregarding the three Ss classified as "others," those Ss who were judged field

dependent on the basis of the traditional Witkin pretest all had ROC functions significantly below the chance diagonal, whereas those initially judged to be field independent obtained ROC functions significantly above. Such an outcome supports the notion that these two groups are distinctly different in ability to detect verticality in this kind of situation.

Judging from the intercorrelations in Table 1, the ROC analysis provided data comparable to those from Witkin's traditional procedure with the Rod and Frame Test. The correlation of ROC area with the Witkin pretest, for example, was almost as high as the reliability of the Witkin measure. Moreover, there was little difference in the size of the correlation coefficients obtained when the ROC area rather than Witkin's measure was used to predict the other measures that Witkin has proposed as tests of field independence or psychological differentiation. These facts lead to the conclusion that Witkin's measure is not contaminated to any great extent by nonperceptual factors in the manner we originally proposed.

The data shown in Fig. 2 indicate that

Table 2  
Analysis of Variance: Discrepancy Scores

Source	SS	df	MS	F
Between Ss	140.46	16		
Dependence (D)	65.00	2	32.50	6.03*
Ss within groups	75.46	14	5.39	
Within Ss	69.39	17		
Frame setting (F)	.05	1	.05	---
DF	22.27	2	11.13	3.31
F x Ss within groups	47.07	14	3.36	

\*  $p < .05$

frame was tilted 28 deg, thus providing further evidence that the dependents differentiated the rod positions only when the frame was tilted 8 deg.

The independents, on the other hand, achieved comparable performance under the 8- and 28-deg frame settings, as indicated by confidence intervals that largely overlap. Yet these equivalent sensitivities were obtained by very different uses of the rating categories in the two conditions.

The sensitivity data for the three Ss in the "others" group is interesting since these Ss displayed, according to the confidence intervals, significant dependence when the frame was tilted 8 deg and significant independence when the frame was tilted 28 deg. Could these Ss represent a transitional stage on the route from dependence to independence? We would prefer not to speculate on this question at present. The data do indicate, however, that these three Ss do represent something more than just a "wastebasket" category. Their individual ROC curves were quite similar for the 8-deg condition and also for the 28-deg condition. Furthermore, their confidence intervals are comparable in range to those of the other

two groups, in spite of the much smaller sample size.

In conclusion, our data do support the notion that Witkin's Rod and Frame Test measures predominantly a sensitivity or perceptual factor. The correlations between the two kinds of measures, those obtained by our method and his, were as high as one might expect to get, even by two different presentations of the Witkin test itself. These data also show that sometimes the perceptual field or context may affect response processes while leaving sensitivity unchanged, as was the case with the independents in this study, whereas at other times, the field or context may change both response processes and sensitivity.

Thus, even though the Rod and Frame task seems far removed from traditional problems in signal detection, the signal detection model has provided evidence that would have been difficult if not impossible to obtain with the traditional Witkin procedure. Moreover, the signal-detection analysis has, in this application, raised some interesting and, we hope, fruitful questions.

#### REFERENCES

GREEN, D. M., & SWETS, J. A. *Signal detec-*

*tion theory and psychophysics*. New York: Wiley, 1966.

GROSS, F. The role of set in perception of the upright. *Journal of Personality*, 1959, 27, 95-103.

HAKE, H. W., & RODWAN, A. S. Perception and recognition. In J. B. Sidowski (Ed.), *Experimental methods and instrumentation in psychology*. New York: McGraw-Hill, 1966. Pp. 331-381.

LINTON, H. Dependence on external influence: Correlates in perception, attitudes, and judgment. *Journal of Abnormal & Social Psychology*, 1955, 51, 502-507.

SHONTZ, W. D. Factors affecting the processing of sequentially presented form parts. Unpublished doctoral dissertation, Iowa State University, 1967.

WITKIN, H. A., DYK, R. B., FATERSON, H. F., GOODENOUGH, D. R., & KARP, S. A. *Psychological differentiation*. New York: Wiley, 1962.

#### NOTES

1. These data were collected by Dr. Cross during his tenure as a NSF Summer (1967) Research Participant at Iowa State University.

2. Address: Bowling Green State University, Bowling Green, Ohio 43402.

(Accepted for publication August 13, 1969.)