
Relationships Between and Reliability Estimates of New (Holtzman) Ink Blot Variables and Conventional Rorschach Scoring Categories

HARL H. YOUNG, University of Oklahoma, Norman

In order to obviate the well-known difficulties of using the Rorschach test as a research instrument, a team of investigators led by Wayne Holtzman (1956a, b) has developed a new ink blot test. Test materials and scoring procedures have been developed which make it possible to measure with adequate reliability such variables as tendency to see movement, primacy of color or shading as a response determinant and form level, while at the same time preserving the rich variety of content, symbolic references, and other qualitative aspects of the Rorschach as it now stands.

The major modifications which have been undertaken consist of increasing the number of inkblots to forty while limiting the number of responses per card to one, and extending the variety of stimulus colors, patterns and shadings used in the original Rorschach materials. The overall goal was to build an instrument which would fullfil the desirable features mentioned and yet possess the psychometric properties necessary for a fruitful research instrument using conventional statistical methods. Preliminary studies (Holtzman 1956a, b; Young, 1957) suggest that the goal has been approximated. At this writing, two alternative forms have been developed, are being published and will be available soon for research purposes.

Modification of scoring procedures consisted of the establishment of variables which in some cases were very similar to conventional Rorschach scoring while in other instances the variables were vastly different. An important question then became apparent. What are the various degrees of relationship between the newly-developed scoring categories and the conventional ones? Obviously, generalizability to the Rorschach variables of the results obtained using the new instrument hinged, in part, upon the answer to this question. The primary objective of this study was to throw light on this particular problem. As a secondary objective, it was planned to obtain internal consistency (content) reliability estimates which would provide a further check on the values obtained in the preliminary studies.

Method

An experimental form of the Holtzman Ink Blot Test was given to 49 males and 48 females, most of whom were summer semester college students. The test was administered in small groups using especially prepared test booklets for recording responses to the colored slides of the ink blots which were projected onto a screen in the front of the room.

Each protocol was scored for the following Rorschach categories using the criteria set forth in Klopfer, *et al* (1954): M, FM, m, W, F, FC, CF, H, Hd, and M:Sum C. Certain Holtzman-type variables were selected and the protocols scored accordingly. These variables will now be described more specifically.

Location (Loc). A score of zero was given each W or whole response; 1 was the score given to a response which utilized a large detail or "D" and a score of 2 was assigned if a "d" or small detail was the area used. Obtained by summing over 40 cards, this score could be called a measure of the subject's tendency to fragment the blots.

Movement Energy Level (Mel). This is a scale adapted from similar measures used earlier by Zubin (1953) and by Wilson (1952). Weights of zero, 1, 2, 3, and 4 were assigned to a given response corresponding to the following categories: zero, no movement; 1, static movement (sitting, facing, lying, etc.); 2, casual movement (walking talking, yawning, etc.); 3, dynamic movement (energetic activity, such as running, jumping, lifting, etc.); 4, violent movement (sobbing madly, whirling, eruptions, explosions etc.).

Form Appropriateness (Fa) and Form Definiteness (Fd). These variables emerged as a result of considering the conventional scoring of form responses as confounding two very different things, the appropriateness of the fit of the percept to the structure of the blot, and definiteness or the specificity of the percept. An example will make this distinction clearer: Responses such as "smoke," "clouds," and "a squashed bug" are very indefinite, yet may be highly appropriate to the blot structure. Similarly, "a Hopi Indian doing a rain dance with the typical rain-making paraphernalia on" is highly definite and specific, yet may be completely inappropriate in terms of the stimulus material. Derived by a slow process involving clinical judgments and the development of scoring manuals, Fa and Fd result from the separation of these two dimensions.

Human Likeness (HL). Another scale developed earlier by Zubin (1953) provided the impetus for the construction of a Human Likeness (HL) Scale. Weights for individual responses were assigned as follows: zero, if a figure was perceived as ennobled or divinized; 1, if the figure was an ordinary human; 2, if the figure

was caricatured or mask-like; and 3, if the figure was a repulsive monster or malformed in some way. This scale was intended to be a measure of the extent to which animate figures are perceived to deviate from the conventional person concept.

Color Scale (*C*). A four-point scoring scheme was used to derive a scale indicating the degree to which a subject used color as a determinant. The scores ranged from zero for a response not using color through 1 for color incidentally used and 2 for color used as a primary determinant, to 3 for color naming and color description.

In addition to the six variables described above, the ratio Mel:C was computed. All the Holtzman variables were correlated with their conventional Rorschach counterparts.

With regard to the reliability estimates, an internal consistency technique based on the matched random subtest method advocated by Gulliksen (1950) was utilized. All ink blots were placed in rank order according to the sum of squares on the variable in question. Then cards were randomly assigned in pairs so that the two halves which resulted would have comparable means and variances. The two halves were then correlated and corrected by means of the Spearman-Brown "prophecy" formula.

Results

The results concerning the relationships between the two scoring procedures are shown in Table I. It is seen at once that, excluding the variables dealing with form level, every value in the table is statistically significant. Moreover, only a few (three) of these values are of a comparatively low order of magnitude. It appears that the Holtzman variables are, in general, highly related to their Rorschach scoring counterparts.

TABLE I

Relations between Holtzman and Rorschach scoring variables*

			Males (N=49)	Females (N=48)
Mel	vs.	M	.79	.91
Mel	vs.	FM	.61	.56
Mel	vs.	m	.36	.29
Loc	vs.	W	—.95	—.95
Fa	vs.	F	—.01	.07
Fd	vs.	F	.12	—.13
C	vs.	FC	.51	.63
C	vs.	CF	.84	.84
HL	vs.	H	.76	.83
HL	vs.	Hd	.35	.51
Mel:M:EC	vs.	M:EC	.73	.87

* .29 required for significance at the .05 level of confidence.

.37 required for significance at the .01 level of confidence.

However, a few exceptions are evident. The form level measures, already mentioned, are the most outstanding in this regard. Since the Fa and Fd dimensions were explicitly designed to separate two components of the usual scoring for form, these results were expected. Indeed, the findings are encouraging in that it seems reasonable to speculate that these scores may be more highly related to such factors as "reality testing" and intelligence than the usual form-level scoring, which has produced disappointingly low correlations in the studies previously reported (Kimbal, 1950; Wilson, 1952).

The correlations for the sexes between Mel and m are also quite low.

Again, this was to be expcted since Mel is scored for any kind of movement while m represents only a small portion of the movement scored under the usual conditions.

Much the same holds for the low correlations between HL and the Hd scores. It is rather difficult for a percept of a part of the body to be construed as divinized or ennobled, while this is not at all unusual for the whole-person response.

Turning now to a consideration of the reliability estimates, the appropriate data are presented in Table II. For the six Holtzman variables, the

TABLE II

Internal Consistency Correlations for the Holtzman Ink Blot Variables

	Subset mean		Subset variance		r_c
Males N=49	X	Y	X	Y	
Loc	7.06	7.10	20.02	19.36	.90
Mel	15.9	16.4	74.04	71.63	.77
Fa	21.4	21.6	14.70	12.90	.67
Fd	42.4	42.2	60.81	88.17	.90
Cl	5.5	5.7	14.41	12.88	.72
Hl	6.6	6.7	17.31	12.7	.68
Females N=48					
Loc	7.9	7.8	18.75	18.24	.91
Mel	16.6	16.0	61.20	53.31	.82
Fa	22.5	22.6	11.08	16.20	.52
Fd	42.4	42.4	69.28	47.12	.77
Cl	5.4	5.5	12.66	13.96	.74
Hl	6.8	6.6	8.94	8.87	.52

reliability estimates range from .67 to .90 among males and from .52 to .91 among females. While some of these values are lower than would be ideally desirable, they are considered adequate for purposes of further research. It is noteworthy that for both sexes, Mel, Loc, and Fd tend to be the most reliable of the six variables.

Synthesizing the foregoing results, it would appear that the Holtzman variables have adequate reliabilities for pursuing validation studies using the new instrument and that, except for the form level variables, m, and Hd, it is fairly clear that the findings will be generalizable in substantial degrees to Rorschach variables. Due to their higher reliabilities, most confidence should be placed in the generalizations involving the Mel and Loc variables.

Discussion

It should be noted that the design of the present study has a feature which might tempt some to temper the conclusions mentioned above. The problem concerns the propriety of applying the Rorschach scoring categories to an instrument different from the one upon which they were originally conceived. In the first place, a control procedure which would have involved applying both scoring systems to the usual Rorschach technique was obviously unfeasible because the inherent properties of the Rorschach measures do not permit adequate reliability estimates. Second, I believe that if Rorschach categories have appreciable validity for measuring personality variables, they should be applicable to almost any well-

constructed ink blot test. At any rate, this study suggests that acceptably reliable measures are highly related to the Rorschach categories and thus may be measuring pretty much the same thing. It is hoped that the ill-rewarded enthusiasm which has characterized Rorschach validation research in the past will now be channelized into the more fruitful leads suggested by these newer methods.

Summary and Conclusions

The Holtzman Ink Blot Test was given to 49 males and 48 females, most of whom were college students. Six newly-developed scoring variables and traditional Rorschach scoring procedures were applied to the protocols. Internal consistency reliability estimates were computed for the Holtzman variables and these measures were then correlated with the Rorschach scoring categories. It was concluded that the Holtzman variables have adequate reliabilities for use in validation research and that, with a few notable exceptions, generalizations based on such studies will be applicable to Rorschach measures to an appreciable extent.

References

Gulliksen, H. 1950. Theory of Mental Tests. New York: Wiley and Sons.

Holtzman, W. H. 1956a. Development of an Experimental Ink Blot Test: A new Departure from the Rorschach (abstract). Am. Psychol. 11: p.-400.

Holtzman, W. H. 1956b. Progress Report of Research on the Development of a new Ink Blot Test (mimeographed). Austin: University of Texas.

Kimbal, A. J. 1950. History of Form Level Appraisal in the Rorschach. J. Proj. Tech. 14: 134-152.

Klopfer, B., *et al.* 1954. Developments in the Rorschach Technique. Vol. 1. Yonkers-on-Hudson: World Book.

Wilson, G. P. 1952. Intellectual Indicators in the Rorschach Test. Unpublished doctoral dissertation. University of Texas, Austin.

Young, H. H. 1957. Personality Test correlates of Orientation to the Vertical: A Test of Witkin's Field Dependency Hypothesis. Unpublished doctoral dissertation, University of Texas, Austin.

Zubin, J. 1953. Experimental Abnormal Psychology. (Preliminary unpublished edition.)