A Creative Personality Scale for the Adjective Check List

Harrison G. Gough
Institute of Personality Assessment and Research
University of California, Berkeley

The Adjective Check List was administered to seven male and five female samples comprising 1,701 subjects. Direct or inferred ratings of creativity were available for all individuals. The samples covered a wide range of ages and kinds of work; criteria of creativity were also varied, including ratings by expert judges, faculty members, personality assessment staff observers, and life history interviewers. The creativity scales of Domino and Schaefer were scored on all protocols, as were Welsh’s A-1, A-2, A-3, and A-4 scales for different combinations of “origence” and “intellectence.” From item analyses a new 30-item Creative Personality Scale was developed. It is positively and significantly ($\rho < .01$) related to all six of the prior measures but surpasses them in its correlations with the criterion evaluations.

Creativity is a valued commodity in every kind of human endeavor. Since the publication of Guilford’s (1950) influential presidential address to the American Psychological Association, an enormous amount of effort has been invested in the study of creativity and its determinants. One line of investigation within the larger domain of inquiry has been the search for methods of assessment that can identify creative talent and potential within the individual. Many of these studies have addressed cognitive issues and problem solving. For example, Guilford and his colleagues (Guilford, Wilson, Christensen, & Lewis, 1951) developed a series of tests stressing ingenuity, the ability to overcome constraining sets, and fluency in ideation. Mednick (1962) proposed a method of assessment requiring the generation of remote associations for the solution of analogies.

In regard to intellectual functioning, it should be noted that most studies have found intellectual ability as usually measured to be unrelated to criteria of originality. MacKinnon and Hall (1972) obtained Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1958) protocols from 88 architects, 37 research scientists, 33 male mathematicians, and 27 female mathematicians who had also been rated on creativity. Within each sample, subjects were dichotomized into those with higher or lower ratings. The higher rated subgroup, taken from all four samples, had a mean full-scale IQ of 133, and the lower rated subgroup had a mean IQ of 131. The difference between the means was not statistically significant, nor were there any noteworthy differences in range or dispersal of scores. In their study of gifted students, Getzels and Jackson (1962) found intelligence to play a smaller role than personality in determining creativity, and Taylor (1960) also found general intellectual ability to be less important than special kinds of thinking and motivational factors.

Ordinary observations have long suggested that artistic temperament and aesthetic dispositions are related to creative potential. A landmark study of this hypothesis was that of Barron and Welsh (1952) in which a nonverbal figure-preference scale was introduced. The original Barron-Welsh Art Scale and the Revised Art Scale (Welsh, 1969, 1975), in which like and dislike responses are balanced, have repeatedly been shown to differentiate between more and less creative persons in various scientific, literary, and artistic fields (Barron, 1972; Welsh, 1977). Another ex-
ample of measurement in the aesthetic domain is the Hall Mosaic Construction Test (Hall, 1972).

Personal traits and dispositions have also been examined in regard to creativity, using both standard personality inventories and specially developed scales and questionnaires. Examples of this kind of work may be found in the writings of Barron (1957, 1958), Domino (1974), Gough (1956/1962, 1976), Helson (1977), Kanner (1976), MacKinnon (1962, 1965), and Stein and Heinze (1960). An overall analysis of intellectual, aesthetic, motivational, and other kinds of tests for creativity has been published by Barron (1965).

A particular topic within the realm of studies of personality is that dealing with the self-concept. The Adjective Check List (ACL; Gough & Heilbrun, 1965) is an assessment device intended for appraising views of the self, and as might be expected it has frequently been employed in investigations of creativity (Cashdan & Welsh, 1966; MacKinnon, 1963; Schaefer, 1969). Several attempts, in fact, have been made to develop creativity scales for the ACL. Schaefer (Smith & Schaefer, 1969) identified 27 items that differentiated between the responses of high school boys rated as more and less creative. Only one of these items (cooperative) was checked more often by those with lower ratings; the other 26 items were more often endorsed by higher rated respondents. Follow-up studies (Schaefer, 1972, 1973) indicated that the scale retained its validity over time.

Domino (1970) asked faculty members of a liberal arts college to identify all male freshmen who had manifested creative ability. Ninety-six students were selected; these students were then matched against 96 unnominated controls on age, IQ, personal adjustment as estimated from the Minnesota Multiphasic Personality Inventory (MMPI), and declared major. In the next year, faculty members were asked to make a special effort to observe the performance of all of these students and to judge which ones had shown evidence of creative ability; this procedure was repeated in the third year. These steps yielded a final sample of 59 creative males and a control sample of 82. For each student, at least three faculty raters completed descriptive ACLs. Composites were formed for the 141 students by assigning a 1 to items checked by two or more raters, and a 0 to items checked by none or by only one observer. An item analysis of the 300 adjectives in the list identified 68 that differentiated between the two subsamples at the .05 level of probability or beyond. Fifty-nine of these were more often used to describe the creative subsample and 9 were used more often to describe the controls. To simplify scoring and analysis, Domino decided to base his scale solely on the 59 items associated with greater creativity. Cross-validation of the scale on the self-report protocols of new samples of males and females produced correlations ranging from .24 to .45 with criterion classifications of creativity.

A third study in which item analyses were made of the ACL was that of Welsh (1975). Two dimensions were first defined, one deriving from intellectual functioning and behavior and termed intellectence, and the other deriving from originality and aesthetic sophistication and termed origence. The interactive grid for these two dimensions permitted the specification of four types of cognitive functioning (Welsh, 1977). Type 1, high on origence but low on intellectence, was characterized by diffuse, global, and imprecise integration with little or no differentiation. Type 2, high on both origence and intellectence, was characterized by synthesis, organization, and the cathexis of metaphor. Type 3, low on both origence and intellectence, was characterized by fragmentation of elements and overattention to details. Type 4, low on origence but high on intellectence, was characterized by analytic and logical preferences. The four types could be very briefly described as imaginative, intuitive, conventional, and analytic in their cognitive styles. Adjective Check List scales were developed for each quadrant by contrasting the responses of individuals in that cell with those of individuals in the other three. The A-1 scale for respondents high on origence but low on intellectence contained 21 items. The A-2 scale for respondents high on both axes contained 25 items. The A-3 and
A-4 scales contained 17 and 24 items, respectively. Scoring on all four scales was by endorsement only; no points were given for non-endorsement of contraindicative adjectives.

During the late 1950s, the writer directed a study of creativity among research scientists (Gough & Woodworth, 1960). When the six ACL scales described above were applied to the 45 scientists in this study, rather disappointing results were obtained. The correlations with the criterion ratings of creativity were .01 for Domino's scale, .08 for Schaefer's, and -.06, -.01, -.08, and .08 for the four measures developed by Welsh. Similar results were obtained when the six scales were scored on the self-report ACLs of the 57 male mathematicians studied by Helson and Crutchfield (1970). Correlations with criterion ratings of creativity were -.11, -.03, .05, -.08, .05, and -.08 for the six ACL scales, in the same sequence. Because of these inconclusive findings, it was decided to undertake a new analysis of the ACL, using larger samples and a broader range of criteria, to see if a stronger measure could be developed.

Method

Samples

Six male samples were available in which the ACL had been administered and criterion ratings of creativity had been gathered. The first was composed of 124 architects studied by MacKinnon and Hall (Hall & MacKinnon, 1969). The second was comprised of the 57 male mathematicians just mentioned (Helson & Crutchfield, 1970). The third included the 45 research scientists also just mentioned (Gough & Woodworth, 1960). The fourth was a previously unreported sample of 530 graduate students in psychology at Berkeley, tested during the first week of entry and rated 3 or 4 years later. The fifth was a sample of 66 seniors in engineering (Gough, 1976). The sixth was a composite sample of 256 men who had been intensively assessed at the Institute of Personality Assessment and Research and rated by staff observers on creativity. This composite sample included 100 Air Force officers (MacKinnon, Crutchfield, Barron, Block, Gough, & Harris, 1958), 70 medical school applicants (Gough & Hall, 1973), 20 college sophomores from an unpublished study of vocational and career planning, 41 males from a study of population psychology (Gough, 1973), and 25 males from a study of environmental preferences (Craik, Note 1).

There were four samples of women for which ACL protocols and ratings of creativity were available. The first was comprised of the 41 women mathematicians studied by Helson (1971). The second included 335 graduate students in psychology at Berkeley, tested at the time of entry and rated between 3 and 4 years later by faculty members. The third was composed of the 51 college seniors reported on earlier by Helson (1967). The fourth was a composite sample of 126 women, including 20 college sophomores from the study of career planning, 41 women from the study of population psychology, 25 women from the study of environmental preferences, and 40 first-year students of law (LaRussa, 1977).

In addition to these 1,631 subjects, there were 35 males and 35 females for whom ACL protocols were available and who had been interviewed by two psychologists and described by them on the ACL and on Block's (1961) 100-item California Q-sort. Although direct ratings of creativity were not obtained, indirect estimates were derived from the ACL and Q-sort portraits furnished by the interviewers, in a way to be described below.

Tests and Criteria

The ACL self-report protocols of all subjects were scored for the creativity scales of Domino and Schaefer and for Welsh's four scales. Four kinds of criterion evaluations were utilized. For the architects, male and female mathematicians, and research scientists, creativity was specified by ratings furnished by expert judges. The validity and reliability of these ratings are discussed in the papers already cited. Ratings by faculty members supplied the criterion for the engineering students, psychology graduate students, and college seniors. The psychology graduate students had entered over a 24-year period. Faculty ratings were obtained every 3 or 4 years, covering students who had entered in the preceding interval. Corrected interjudge reliabilities for these ratings of creativity ranged from .33 to .87, with a median of .77. Equivalent coefficients were found for the engineers and college seniors. The 256 males and 126 females seen at the institute in intensive programs of assessment were rated on creativity by panels of 10 or more observers. Corrected interjudge reliabilities for these ratings were typically very high, ranging from .80 to .98.

To develop a criterion for the 70 interviewed subjects, the interviewers' checks on five adjectives and placement of five Q-sort items were summed. The five adjectives were imaginative, insightful, intelligent, original, and resourceful. There were four Q-sort items given positive weighting: Has a wide range of interests, Appears to have a high degree of intellectual capacity, Thinks and associates to ideas in unusual ways; has unconventional thought processes, and Able to see to the heart of important problems. One Q-sort item was assigned a negative weight: Is uncomfortable with uncertainty and complexities. The total score based on these 10 components had a standardized item alpha reliability coefficient of .79.
Analyses

For purposes of item analysis, four subgroups were defined: (a) 558 males from the samples of architects, mathematicians, scientists, engineers, and other assessed males; (b) 530 male graduate students in psychology; (c) 218 females from the samples of mathematicians, college seniors, and other assessed females; and (d) 335 psychology graduate students. Ratings of creativity were converted to standard scores by subsample. Point-biserial coefficients of correlation were calculated between each of the 300 items in the ACL and the criterion standard scores.

Results

Thirty items were selected for inclusion in the Creative Personality Scale (CPS) on the basis of item analysis. Several examples may be given. The item egotistical yielded correlations with the criterion ratings of creativity of .17, .06, .08, and .08 in the four subgroups. Although only the first of these coefficients is statistically significant ($p < .01$), all are positive, and the item itself is consonant with prior conceptualizations of the creative personality. The item also appears on Domino's scale. The item original produced correlations of .13, .06, .17, and .11. Two of these coefficients are statistically significant ($p < .05$), and the word appears on both the Schaefer and Domino scales. It is also quite clearly consonant with prior conceptualizations of the creative personality.

The item conservative had correlations with the criterion of -.16, -.08, -.13, and -.10. Two of these coefficients are significant at the .05 level of probability, and one is significant at $p = .06$. Because of these findings and because absence of the attribute appears to be consonant with previous conceptualizations of the creative personality, the item was retained with a negative weighting. In this way, 18 positive and 12 negative items were selected for the final scale. The positive items were capable, clever, confident, egotistical, humorous, individualistic, informal, insightful, intelligent, interests wide, inventive, original, reflective, resourceful, self-confident, sexy, snobbish, and unconventional. The negatively weighted items were affective, cautious, commonplace, conservative, conventional, dissatisfied, honest, interests narrow, mannerly, sincerer, submissive, and suspicious. In scoring a protocol, 1 point is given each time one of the 18 positive items is checked, and 1 point is subtracted each time one of the 12 negative items is checked. The theoretical range of scores is therefore from $-12$ to $+18$.

The 30 items were then compared with those in the previous six scales. Seven CPS items were found in Schaefer's 27-item scale, all scored in the same direction. Fifteen CPS items were found in Domino's 59-item scale; 14 were scored in the same direction, the exception being dissatisfied, which received a negative weight on CPS and a positive weight on Domino's measure. Two CPS items were found to overlap with Welsh's A-1 scale, 5 with his A-2 scale, 1 with his A-3 scale, and 2 with his A-4 measure. Except for dissatisfied on A-2, all of these common items were scored in the same direction. Because dissatisfied showed these two reversals, attention was redirected to the item analytic data. On the four subgroups the item had correlations of -.02, -.07, -.10, and -.02. Additional analyses were then computed on the male mathematicians, the research scientists, and the college seniors, where the coefficients were -.17, -.36, and -.10. Because of these findings it was decided to retain the item with a negative scoring weight.

The 30-item Creative Personality Scale (CPS) was then scored on all of the samples included in the present study. Alpha coefficient reliabilities were computed on the four subgroups defined for the item analysis. The coefficients were .77 for the male composite group, .73 for the male graduate students, .81 for the female composite group, and .73 for the female graduate students. Means and standard deviations for each sample are given in Table 1. Among the male samples, the highest mean was attained by the research scientists, followed closely by the psychology graduate students. The lowest mean was that for the 35 males in the sample seen in interviews only. Among the females, the highest mean was that for the female graduate students, and the lowest was that for the women seen only in interviews. Where comparisons seem appropriate, as between the two samples of graduate students and the two samples of
interviewed subjects, CPS scores were significantly \( p < .05 \), albeit only slightly, higher for males than for females.\(^1\)

Table 2 gives the correlations among the seven ACL scales, computed on the samples of male and female graduate students in psychology. The median coefficients between CPS and each of the other measures were .68 for Domino’s scale for creativity, .74 for Schaefer’s scale, .42 Welsh’s A-1, .51 for Welsh’s A-2, .32 for Welsh’s A-3, and .32 for Welsh’s A-4. For the other measures, the highest median correlations were those of .89 between the Domino and Schaefer scales, .64 between A-1 and the Domino and Schaefer scales, .82 between A-2 and the Domino and Schaefer scales, .64 between A-3 and A-4, and .64 between A-4 and the Domino and Schaefer scales. The presence of the 12 negatively weighted items in CPS appears to be the reason why the other six scales correlate more highly among themselves than they do with CPS.

Table 3 gives the correlations of the seven ACL scales with criterion ratings for creativity. For all seven of the male samples, the highest coefficient each time was that for CPS, and in six of the seven instances the coefficients for CPS were significant at or beyond the .05 level of probability. The exception was

<table>
<thead>
<tr>
<th>Scale</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Domino</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>creativity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>.89</td>
<td>.64</td>
<td>.82</td>
<td>.51</td>
<td>.63</td>
<td>.69</td>
</tr>
<tr>
<td>Female</td>
<td>.89</td>
<td>.64</td>
<td>.81</td>
<td>.54</td>
<td>.65</td>
<td>.67</td>
</tr>
<tr>
<td>2. Schaefer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>creativity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>—</td>
<td>.63</td>
<td>.76</td>
<td>.47</td>
<td>.55</td>
<td>.73</td>
</tr>
<tr>
<td>Female</td>
<td>—</td>
<td>.61</td>
<td>.73</td>
<td>.46</td>
<td>.48</td>
<td>.76</td>
</tr>
<tr>
<td>3. Welsh A-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>—</td>
<td>.62</td>
<td>.51</td>
<td>.44</td>
<td>.45</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>—</td>
<td>.62</td>
<td>.46</td>
<td>.35</td>
<td>.40</td>
<td></td>
</tr>
<tr>
<td>4. Welsh A-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>—</td>
<td>.31</td>
<td>.43</td>
<td>.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>—</td>
<td>.36</td>
<td>.42</td>
<td>.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Welsh A-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>—</td>
<td>.69</td>
<td>.32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>—</td>
<td>.61</td>
<td>.33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Welsh A-4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>—</td>
<td>.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>—</td>
<td>.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Creative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. All coefficients are statistically significant beyond the .01 level of probability. A-1 = Welsh high origence, low intellectence; A-2 = Welsh high origence, high intellectence; A-3 = Welsh low origence, low intellectence; A-4 = Welsh low origence, high intellectence.

The coefficient of .25 for the 45 research scientists, where \( p = .097 \) in a two-tailed test. The smallest coefficient for CPS was that of .15 for the psychology graduate students, and the largest was that of .42 for the 256 males rated by observers in the various assessments. On

\(^1\) For purely normative purposes, Adjective Check List protocols for 1,121 college females and 760 college males from the author's research files were scored for the new 30-item scale. The mean of 5.03 \((SD = 4.01)\) for males was significantly higher \((p < .01)\) than that of 3.97 \((SD = 4.34)\) for females. This difference suggests that where raw scores on the new scale are used, analyses should be conducted separately for males and females. It should be noted that most studies with the ACL use scale scores that have been standardized by sex.
The first six male samples, higher coefficients for CPS would be expected because these samples were used in the item analyses. The 35 interviewed males, however, were not used in the item studies and therefore constitute a true cross-validating sample. In this sample the only statistically significant coefficient was that of .35 for CPS.

For the five samples of women, the coefficient for CPS was the highest in four instances and statistically significant at the .05 level each time. The exception occurred in the sample of 41 women mathematicians. For these subjects, Welsh's A-3 scale had a statistically significant (p < .01) coefficient of -.42, as did his A-4 scale, for which the coefficient was -.40. The coefficient for CPS on this sample was .28, p = .075 in a two-tailed test. The last sample reported in Table 3, the 35 women seen only in life history interviews, constituted a true cross-validating sample. For these subjects, there were three statistically significant (p < .05) correlations between scales and the derived criterion rating of creativity: .34 for Schaefer's scale, .35 for Welsh's A-4 scale, and .40 for CPS.

In the six male samples from which item-analytic data were drawn, the median coefficient between CPS and the criterion ratings of creativity was .30, and in the small cross-validating sample of 35 males, the correlation between CPS and the inferred rating of creativity was .35. A similar summary for the female samples yielded a coefficient of .28 for the samples used in the item analyses and a coefficient of .40 for the small cross-validating sample. These samples covered a wide range of ages, kinds of work, and circumstances of testing. They also involved the use of four vantage points in rating creativity: expert judges, faculty members, personality-assessment staff observers, and life-history interviewers. If one-tailed tests of significance are allowed, the new 30-item scale can claim a
statistically significant relationship with every criterion in every sample. If the two-tailed test is used, 10 of the 12 validity coefficients for CPS were significant at the .05 level of probability or beyond. On the basis of these findings, it seems reasonable to conclude that the new scale is a reliable and moderately valid measure of creative potential and that it may properly be included among the scales to be scored on the Adjective Check List.

Reference Note

1. Craik, K. H. Impression of a place: Effects of media, context, and personality. In S. Saegert (Chair), Psychology of the Urban Environment. Symposium presented at the annual meeting of the American Psychological Association, Toronto, August 1978. (Copies available from the Institute of Personality Assessment and Research, University of California, Berkeley, California 94720.)

References


Barron, F. The psychology of imagination. Scientific American, 1958, 199, 150-156.


 Guilford, J. P., Wilson, R. C., Christensen, P. R., & Lewis, D. J. A factor analytic study of creative thinking: I. Hypotheses and descriptions of tests. Reports from the Psychological Laboratory (No. 4). Los Angeles: University of Southern California, 1951.


LaRussa, G. W. Portia’s decision: Women’s motives for studying law and their later career satisfaction as attorneys. Psychology of Women Quarterly, 1977, 1, 350-364.


Schaefer, C. E. Follow-up study of a creativity scale for the Adjective Check List. Psychological Reports, 1972, 30, 662.


Received October 3, 1978