MEASURING LEVELS OF EMOTIONAL DEVELOPMENT*

Department of Psychology, University of Wisconsin-Madison;
and School of Education, Northwestern University

DAVID F. GAGE, PHILIP A. MORSE, AND MICHAEL M. PIECHOWSKI

Acknowledgments .................................................. 130
Summary .......................................................... 131
I. Introduction .................................................... 132
II. Method ........................................................ 134
III. Results ......................................................... 137
IV. Discussion ..................................................... 145
V. Conclusion ....................................................... 149
Appendices ........................................................ 149
References ......................................................... 152

* Received in the Editorial Office, Provincetown, Massachusetts, on February 5, 1980.
ACKNOWLEDGMENTS

Portions of this paper are based on the senior author's undergraduate thesis submitted to the Department of Psychology, University of Wisconsin-Madison. We are particularly indebted to the eight students who volunteered their time and enthusiasm for this study and to Dr. Roberta Ray for her invaluable assistance in facilitating our contacts with the students. In addition, we are grateful to Dr. John Townsend of the Department of Psychology, Catholic University of America, for his many helpful suggestions and encouragement, especially his much appreciated statistical advice. Thanks are extended to Alison Baker for her creative suggestions regarding the development of the situations and the questionnaire items. Finally, special thanks are due to Kathleen Piper for her skillful rating, her many constructive criticisms, and her sustaining enthusiasm and involvement in this work. Reprint requests should be addressed to the first author at the first address shown below.

DAVID F. GAGE
PHILIP A. MORSE*
MICHAEL M. PIECHOWSKI**

Department of Psychology
The Catholic University of America
Washington, D.C. 20064

* Department of Psychology
University of Wisconsin-Madison
Madison, Wisconsin 53706

** School of Education
Northwestern University
Evanston, Illinois 60201

SUMMARY

According to the theory of positive disintegration there are distinct levels of emotional and personality development that are not tied to specific chronological ages. The present study sought (a) to develop a set of methods for assessing a person's developmental level, and (b) to assess the convergent and discriminant validity of the methods. Over a period of one month, eight college seniors wrote autobiographies and completed four other measures. Three of the four new methods exhibited consistent convergent and discriminant validity. The relative advantages and disadvantages of the various methods are discussed.
I. INTRODUCTION

This paper describes a preliminary attempt at developing a set of methods for assessing levels of development as defined by Dabrowski’s theory of positive disintegration (4, 5, 6, 8). Detailed description of levels and prior research is given elsewhere (8, 14, 15) and need not be repeated here.

The first purpose of this study was to make future research employing the theory of positive disintegration easier by developing methods that could assess people’s levels of psychological functioning more efficiently and economically than the autobiographical and clinical approaches previously employed. In the present study three tests of increasing objectivity were developed for assessing expressions of the underlying themes of six dynamisms.

The second purpose was to assess the construct validity by comparing the new methods to each other and to the autobiography method with the use of the convergent-discriminant methodology of Campbell and Fiske (3). This methodology was chosen as the most appropriate approach, since adequate criterion data are difficult, if not impossible to obtain (1).

The third purpose was to determine whether the assessment of a person’s level of development could be made easier by expanding Dabrowski’s concept of a dynamism to include expressions which might also be observed at lower levels. Dabrowski described 30 “dynamisms” (e.g., Inferiority toward Oneself, Dissatisfaction with Oneself) as the elements which constitute the developmental structures called levels. But because 75 percent of the dynamisms are centered at Level III or higher, Levels I and II are difficult to assess. In fact, “Level I is defined by the total absence of dynamisms” (7, p. 27). This creates a serious problem because the absence of a variable is difficult to quantify. The present study tested whether or not it is feasible to resolve this problem by expanding the concept of a dynamism. To accomplish this, the underlying themes of six of the 30 dynamisms were defined. Then the expressions of the themes of each of the levels were also provided (see Appendix A for theme definitions; Appendix C illustrates the different expressions of one of the themes employed in the Situation-choice method described below).

In general, the present research examined six themes (Susceptibility to the Influence of Others, Personal Conflict, Inferiority, Dissatisfaction, Self-observation, and Personality Ideal) across three levels and two intermediate levels, using five different methods: Autobiography, Definition-response, Situation-choice, Situation-reason, and Questionnaire.

The three purposes of this research, then, were to (a) develop new assessment methods, (b) assess construct validity, and (c) determine whether assessing a person’s developmental level can be simplified by measuring expressions instead of dynamisms.

---

1 One of the reasons for the emphasis on the higher levels is that Dabrowski’s clinical work largely focused on people undergoing accelerated personality growth in psychotherapy who were primarily at the higher levels.

2 The term “expression” was used here to denote a specific manifestation of a theme at one of the levels or borderline of two levels. Whereas dynamisms are considered to operate only at certain levels (cf. 14, Figure 1), a theme is operationally defined as having expressions at all levels. For example, the dynamism, Inferiority toward Oneself, found only above Level II, describes a debasement of oneself as a result of the perception that one is not acting or changing in harmony with one’s developing hierarchy of values. As presently defined, the expression of Inferiority at Level I was a response of indignation or externalized blame emanating from a private belief that one does not possess certain physical comforts, social privileges, material amenities, etc., to which one feels entitled. The Level II expression is a self-demeaning response to perceiving oneself as not being as intelligent, popular, proficient, etc., as someone else, or not as close to the “societal ideal” as one would like to be. The five other themes presented in Table 1 were similarly operationalized into their component expressions.
II. METHOD

A. Subjects

The Ss were eight college seniors enrolled in an introductory clinical psychology course. The six female and two male students ranged in age from 20 to 27. Participation was voluntary and fulfilled part of the course requirements.

B. Materials

The five methods described below were used to assess the presence of the six themes listed in Table 1.

The Autobiography (A) method has been the method most used in studies of the theory of positive disintegration (7, 8). The written instructions used in the present project were slightly modified and asked Ss “... to focus on those events, feelings, insights—in fact any type of occurrence, natural or spiritual—which had, or may be having right now, an impact on your development ...” Ss were requested to write at least 10 pages and to write them in more than one sitting. Responses in the autobiographies which expressed one of the six themes were identified and rated for level: I, I-II, II, etc.

The Definition-response (D) method was designed to focus the Ss’ attention on each theme without giving any obvious cues to particular expressions of the theme, and to allow the Ss to respond in an open-ended, free-response format similar to the A method. Six short descriptive statements presented the six themes (see Appendix B). Each statement was written on a separate page (without the name of the theme) along with instructions requesting the Ss to describe up to six personal experiences corresponding, as closely as possible, to how they experienced the theme. Furthermore, the Ss were informed that on the day after making the lists, they should add pertinent experiences they may have subsequently remembered. When each list was completed, they were asked to rank the items in the list according to how frequently or how strongly they felt that experience. The items in the lists were rated with regard to the level of expression.

The Situation-choice (C) method was modelled after Kohlberg’s (10) moral dilemmas. It attempts to increase the objectivity of both the stimulus and the response by presenting the same paragraph-long situation and 12 one-sentence choices to every S. The situations were 12 short paragraphs, each describing a plausible real-life dilemma, and each indicative of a particular theme (two situations per theme). Each situation concluded with a question related to the main idea of the paragraph, and 12 alternative choices of possible response to the situation. Six of the responses were worded as different reasons for answering “yes,” and six for answering “no.” Each reason was characteristic of one of the three primary levels. Four choices (two “yes” and two “no”) incorporated Level I reasoning, four choices used Level II reasoning, and four used Level III reasoning. The authors worked independently to determine the level designations of the alternatives and employed only those on which they all agreed. As an example, a situation depicting the theme, Susceptibility to the Influence of Others, with its 12 alternatives is presented in Appendix C. The instructions were to rank the four choices that were closest to how the S “would most likely respond in that situation,” and to include at least one “yes” and one “no” response among the four.

The Situation-reason (R) method was developed because the alternatives a person selects may not always be perfectly related to his or her reason for making the specific choice (2, 16). This method asked the Ss to explain their rationales for selecting specific choices—i.e., how they perceived the situation, or what meaning it had for them. The responses were rated according to levels. The situations were the same as in the C method, but the specific stimuli were the S’s first two choices. The Ss were to reread their first two choices for each situation and then write their reason for selecting each choice.

The Questionnaire (Q) approach was included in this study to determine if a highly objective assessment device would achieve a satisfactory degree of convergent validity with the other measures. There were 72 one-sentence items on the questionnaire. Each item represented one of the themes (12 items per theme) at one of the three levels (24 items per level). For example, Personality Ideal has four items (two stated in the positive direction and two in the negative) at each of the three levels. Items were selected which received the same theme and level ratings by the authors working independently. Examples of the items are, “I enjoy competing against others only in those events or games in which I know I can win” (Susceptibility to the influence of others-Level I), “I see myself as many different types of people, and none seems clearly better than the others” (Self-observation-Level II), “I would feel that I, myself, was inadequate if a friend came to me for help and he or she wasn’t helped” (Inferiority-Level III). The instructions were to rank each item on a six-point scale ranging from “strongly agree” to “strongly disagree.”
C. Procedure

The five different measures were presented in the order described above. In each set of instructions students were briefly encouraged to respond openly and honestly. They were assured that their responses would in no way affect their grades in the course. To help assure anonymity, the measures were distributed and collected as a group and the students used a code number instead of their names. One week was allotted for each of the methods with the exception of the C and R methods which were completed in the same week.

III. Results

A. Interrater Reliability

All reliabilities were obtained by correlating the agreement of the raters on each individual response rather than by simple correlating overall agreement. In other words, if one rater rated five consecutive items "a, b, c, c, e" and the second rated them "a, c, c, b, e," the overall agreement is 1.0, but on an item-by-item correlation it is .6. We have used the latter method.

1. Autobiography

The first author determined what constituted a response unit (14, p. 270), then he and a trained rater independently scored the autobiographies with respect to the theme and level of expression of the responses (six themes times five levels). Agreement on a response was registered only when both theme and level were identically rated. Rating by this method was extremely complex because of the large number of possible combinations. The interrater reliability was .66.

2. Definition-Response

All of the Ss' responses were retyped randomly in lists by theme to insure the rating of a given response was not influenced by prior ratings. This method stipulated the theme so only level was scored. The interrater reliability for this method was .80.

3. Situation-Reason

The Reasons responses were also combined and mixed across Ss by situation. Also, to insure more complete independence between the R and C methods, the reasons were rated without knowledge of the choices with which they had been associated. The interrater reliability was .75.

B. Data Analysis

On each method Ss were given scores for the different levels which were the percentages of their responses at the different levels. For example, each S has a score for demilevel II-III on the R method which is the ratio of his or her number of II-III responses to his or her total number of responses on R. Percentages were used because the number of response units varied across methods and across Ss.\(^8\) With the use of the percentage scores, Ss were

\(^8\) As a check on the assumption that the total number of responses is not related to the level of functioning (7, 8), a correlation was computed on these two dependent variables. The relation-
ranked on each method at each level. Spearman rank-order correlation coefficients were used to establish the degree and direction of the relationships among the rankings. The correlations between all possible pairs of levels and methods are presented in Table 1.

The construct validity was assessed by means of the convergent and discriminant criteria outlined by Campbell and Fiske (3). In their "multitrait-multimethod" analysis (here relabeled "multilevel-multimethod"), four classes of correlations are possible: (a) monomethod-monomethod (these correlations represent test-retest, split-half, or alternative-form reliabilities, but were not obtained in the present study due to limits on the S's time), (b) heteromethod-monomethod (convergent validity), (c) monomethod-heterolevel (discriminant validity), and (d) heteromethod-heterolevel (also discriminant validity).

C. CONVERGENT VALIDITY

In order to establish convergent validity, different methods of measuring the same levels should show significant correlations (see the first five columns in Table 1). Because of the small N only correlations greater than .63 are significant ($p < .05$). Accordingly, seven coefficients (ranging from .67 to .96) are reliable. However, 16 correlations are sufficiently large ($r > .45$) to account for more than 20 percent of the variance.

The absolute number of significant convergent correlations among the different methods is greatly restricted by the small sample size. Also, the eight Ss within the sample are remarkably close in their measured levels of development. All of them fall primarily in the II to II-III range. A larger spread in measured levels could be expected to produce higher reliability and consequently greater validity as well. Because of the preliminary nature of this study no effort was made to obtain a more heterogeneous sample.

Inspection of the monomethod coefficients of Table 1 reveals that the Q method consistently lacks convergent validity with any of the other methods; thus it was not measuring what the other tests measured. All but one of the negative correlations appearing in these columns is associated with the Q method, and none of these negative correlations approaches significance.

Assessments of Level II show by far the greatest convergence, with the two lowest correlations on the A, D, C, and R methods being .58. Further, with the exception of one correlation between the C and D methods (and all with the Q method) the Level III coefficients are all high. Finally, all Level I convergent values are moderate or high. Although monolevel values significantly exceed heterolevel values, there are certain exceptions. A few high correlations exist between I-II and the adjacent Levels I and II which may challenge the usefulness of maintaining a distinction between I-II and its adjacent levels. However, if differential relationships with other levels can be observed, then the levels can be considered empirically distinct (12). For example, moderate correlations are noticeable between Levels I and II, but their relationships to other levels differ. Level I exhibits slightly negative relationships with II-III, whereas II and II-III exhibit more highly negative relationships (the difference between the two is significant, $p < .05$, Wilcoxon matched-pairs, signed-ranks test, two-tailed). Similarly, I and III are negatively related, but II and III are much more negatively related ($p < .05$). Such differential relationships support the contention that the two levels are distinct even though they are moderately correlated. Similar distinctions are not observed between I-II and its adjacent levels or between II-III and III.

D. DISCRIMINANT VALIDITY

While convergent validity demands that high correlations be obtained between different measures of the same level, discriminant validity requires (a) the convergent correlations be larger than correlations between different levels with the use of different measures, (b) the convergent correlations also be larger than correlations between different levels with the use of the same measures, and (c) the same pattern of level interrelationships exists among the different levels found in the rows of Table 1 (e.g., certain levels which correlate more highly than others must correlate more highly regardless of the measures used to obtain the correlation).

The first requirement is assessed by summarizing approximately 200 comparisons between convergent validity values and their corresponding heteromethod-heterolevel values (see Table 2). For the A, D, C, and R methods, the convergent values are consistently greater than the heterolevel values ($p < .001$, two-tailed sign test). In the typical validation study, personality constructs are ideally independent, so convergent values must be greater than heterotrait values in an absolute sense. However, in this case negative heterolevel values are predicted, so magnitude and direction were considered. Campbell and Fiske (3, p. 98) employed the sign test as a rough estimate of consistency while noting the fact that different comparisons were probably not independent of one another.

ship was negligible ($r = .02$) for the A method which exhibited the greatest variability in total number of responses.

4 Specifically, the developmental levels for each of the eight Ss are as follows: 2.0, 2.0, 2.2, 2.2, 2.3, 2.3, 2.4, 2.5. These values are mean level scores averaged across all six themes and the four methods which evidenced good convergent validity (A, D, C, and R.).
### Table 1: Correlations between Levels as Measured by the Same and by Different Methods

<table>
<thead>
<tr>
<th></th>
<th>Monolevela</th>
<th>Distance between levels</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
<td>I-II</td>
<td>II-III</td>
<td></td>
<td>I-I-II</td>
<td>II-III</td>
<td></td>
<td>I-I-II</td>
<td>II-III</td>
<td></td>
<td>I-I-II</td>
</tr>
<tr>
<td>Methods</td>
<td>I</td>
<td>I-II</td>
<td>II-III</td>
<td></td>
<td>I-I-II</td>
<td>II-III</td>
<td></td>
<td>I-I-II</td>
<td>II-III</td>
<td></td>
<td>I-I-II</td>
</tr>
<tr>
<td>AA</td>
<td>13</td>
<td>63</td>
<td>-57</td>
<td>42</td>
<td>-51</td>
<td>-66*</td>
<td>-26</td>
<td>-51</td>
<td>-66*</td>
<td>-26</td>
<td>-51</td>
</tr>
<tr>
<td>DD</td>
<td>21</td>
<td>-05</td>
<td>-79*</td>
<td>35</td>
<td>-51</td>
<td>-66*</td>
<td>-26</td>
<td>-51</td>
<td>-66*</td>
<td>-26</td>
<td>-51</td>
</tr>
<tr>
<td>GC</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>RE</td>
<td>83*</td>
<td>54</td>
<td>-72*</td>
<td>65*</td>
<td>-51</td>
<td>-66*</td>
<td>-26</td>
<td>-51</td>
<td>-66*</td>
<td>-26</td>
<td>-51</td>
</tr>
<tr>
<td>QQ</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AD</td>
<td>38</td>
<td>19</td>
<td>68*</td>
<td>47</td>
<td>61</td>
<td>-51</td>
<td>-08</td>
<td>-45</td>
<td>-62</td>
<td>27</td>
<td>-36</td>
</tr>
<tr>
<td>AC</td>
<td>24</td>
<td>-62</td>
<td>-51</td>
<td>51</td>
<td>-71*</td>
<td>-52</td>
<td>-03</td>
<td>-33</td>
<td>-34</td>
<td>-34</td>
<td>-65</td>
</tr>
<tr>
<td>DA</td>
<td>38</td>
<td>19</td>
<td>68*</td>
<td>47</td>
<td>61</td>
<td>94*</td>
<td>-16</td>
<td>40</td>
<td>64*</td>
<td>-77*</td>
<td>-65*</td>
</tr>
<tr>
<td>DC</td>
<td>25</td>
<td>-84*</td>
<td>-13</td>
<td>-23</td>
<td>-40</td>
<td>-10</td>
<td>-03</td>
<td>-33</td>
<td>-34</td>
<td>-34</td>
<td>-65</td>
</tr>
<tr>
<td>DR</td>
<td>55</td>
<td>42</td>
<td>67*</td>
<td>11</td>
<td>58</td>
<td>-27</td>
<td>-52</td>
<td>30</td>
<td>70</td>
<td>17</td>
<td>-52</td>
</tr>
<tr>
<td>DQ</td>
<td>-09</td>
<td>-05</td>
<td>-33</td>
<td>-13</td>
<td>-48</td>
<td>-30</td>
<td>-20</td>
<td>-03</td>
<td>51</td>
<td>-03</td>
<td>51</td>
</tr>
<tr>
<td>RA</td>
<td>24</td>
<td>-62</td>
<td>-51</td>
<td>14</td>
<td>-07</td>
<td>21</td>
<td>-69*</td>
<td>-04</td>
<td>-38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RD</td>
<td>25</td>
<td>-84*</td>
<td>-13</td>
<td>-23</td>
<td>-40</td>
<td>-10</td>
<td>-03</td>
<td>-33</td>
<td>-34</td>
<td>-34</td>
<td>-65</td>
</tr>
<tr>
<td>RC</td>
<td>72*</td>
<td>-96*</td>
<td>-75*</td>
<td>37</td>
<td>-75*</td>
<td>-08</td>
<td>-75*</td>
<td>-08</td>
<td>-46</td>
<td>-46</td>
<td>-46</td>
</tr>
<tr>
<td>QA</td>
<td>07</td>
<td>-19</td>
<td>-04</td>
<td>-02</td>
<td>-20</td>
<td>-54</td>
<td>-24</td>
<td>22</td>
<td>-24</td>
<td>22</td>
<td>-24</td>
</tr>
<tr>
<td>QD</td>
<td>-09</td>
<td>-05</td>
<td>-33</td>
<td>14</td>
<td>-02</td>
<td>-30</td>
<td>-01</td>
<td>15</td>
<td>-01</td>
<td>15</td>
<td>-01</td>
</tr>
<tr>
<td>QR</td>
<td>-23</td>
<td>-24</td>
<td>-11</td>
<td>-19</td>
<td>-43</td>
<td>-08</td>
<td>-21</td>
<td>-11</td>
<td>-11</td>
<td>-11</td>
<td>-11</td>
</tr>
<tr>
<td>Column average</td>
<td>44</td>
<td>63</td>
<td>73</td>
<td>18</td>
<td>53</td>
<td>40</td>
<td>38</td>
<td>41</td>
<td>30</td>
<td>23</td>
<td>41</td>
</tr>
<tr>
<td>Average for a given distance</td>
<td>50</td>
<td>16</td>
<td>-24</td>
<td>-23</td>
<td>-37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: For example, the correlation $A_1 D_{50} (0.67)$ means that the ranking of the $S$'s Level I scores on the A method correlated $0.67$ with the ranking for Level I-II on the D method. This should not be confused with the correlation $D_{50} A_{50} (0.94)$ which corresponds to the ranking of $I$ on the D method and I-II on the A method.

The Q method was not included in the averages for either the columns or the specific distances between levels.

* In the monolevel columns $A_1 D_1 = D_1 A_1$. Each coefficient is presented twice.

* $p < .05$, one-tailed.
coefficients are not observed to be consistently greater. The rows in Table 2 show that the primary levels have fewer exceptions than the borderline levels. Seventy percent of the exceptions listed in the table are due to correlations of a level with an adjacent level exceeding the validity value.

The second requirement is that a level should correlate higher with the same level measured by a different instrument (e.g., the correlation $A_1D_1 = .38$) than with different levels measured by the same instrument ($A_1A_{11} = .24, A_1A_{111} = -.51$, and $A_1A_{111} = -.26$). Numerous exceptions to this requirement would mean the validity values were inflated, probably as a result of method variance. The summary of comparisons in Table 3 reveals that the exceptions by either method or level are far fewer than would be expected by chance.

The third requirement states that if the levels are assessed accurately by the various methods, then the pattern of correlations among the levels will be similar whichever methods are employed, and the methods may be considered interchangeable. The interrelationships of interest are the heterolevel columns of Table 1 (i.e., all but the first five columns). In order to determine whether any pattern actually exists, the columns are ranked from 1 to 10 across each row (by method). The C and Q methods are omitted because they do not provide a sufficient number of interrelations for the analysis, leaving nine rows to be intercorrelated. Among the nine rows, 37 correlations are obtained, 23 of which reach significance despite the small number of ranks ($p < .05$, Spearman rank correlations, one-tailed test).

E. RELATIONSHIPS BETWEEN LEVELS

Reasoning from the hierarchical order of levels postulated by the theory of positive disintegration, it is possible to predict that any two adjacent levels will be more positively correlated than levels farther apart. Inspection of the averages in the bottom row of Table 1 reveals that same level correlations are significantly higher than those one-half level apart ($p < .05$; Wilcoxon matched-pair signed-ranks test, one-tailed). This result is obtained by analyzing pairs of coefficients that are based on the same two methods but differ in the level of one member of the pair (e.g., $D_{11}D_{11}$ with $D_{11}D_{11}$). Coefficients a half-level distant are higher than those one level apart ($p < .01$); for example, $D_{11}D_{11}$ is paired with $D_{11}D_{11}$.

Furthermore, with reference to correlations between levels greater than a half increment apart (e.g., zero vs. one level distant, or half vs. two levels distant), the result in five out of six cases is that the greater the distance apart, the more negative the relation ($p < .01$). The only exceptions is that the correlations between levels I and III are not significantly more negative than those one level apart.

Another way to assess the relationships among the levels is to specify which heterolevel correlations exceed the convergent validity coefficients. Examination of Table 1 reveals that when heterolevel values exceed the
monolevel values it is most often a correlation involving a level only a half step away. Half-level correlations account for 70% of the higher correlations, one-level differences for about 20%, one and a half-level differences for 7%, and a two-level difference occurs only once (3%). For example, the demilevel I-II is sometimes significantly positively correlated with Levels I and II, and in a couple of instances the demilevel II-III is highly correlated with Level III. Thus, the pattern of interlevel correlations obtained in this study is generally consistent with a sequential/hierarchical ordering of the levels in the theory of positive disintegration.

IV. DISCUSSION

The results show that each of the three purposes of this study was accomplished. Regarding the first two purposes, three of the four methods which were developed to substitute for the Autobiography method evidence both convergent and discriminant validity. The D, C, and R methods converge, or agree, upon what is being measured, while the Q method demonstrates no sign of such agreement: thus the benefit gained from increases in objectivity in the D, C, and R methods disappears for the more objective Q method.

The D, C, and R methods quite conclusively satisfy the requirements for establishing discriminant validity. First, convergent values are consistently higher than correlations of different levels on different tests. Second, these values are higher than correlations of different levels on the same test. This second observation points to the fact that common-level variance is greater than common-method variance, an important finding given that all the devices are basically paper-and-pencil, self-report measures. Although the convergent values significantly exceed the heterolevel values, exceptions were noted. Possibly they raise some question as to whether or not the intermediate levels can be meaningfully differentiated from primary levels. Another possibility, however, is that the relationships for the intermediate levels are actually less reliable than those for the full levels. Supporting this hypothesis is the observation that proportionately more disagreements occurred in intermediate-level ratings than full level ratings. This question needs to be addressed more directly in future research. The third, and final, discriminant validity requirement is satisfied because the pattern of interrelationships among the levels is consistent across the A, D, C, and R methods, arguing persuasively for the interchangeability of the methods (9).

The results demonstrate that the third purpose of the study was accomplished: namely, that assessing a level of emotional development as defined by the theory of positive disintegration is simplified by expanding the theory's concept of a dynamism to that of a theme with characteristic expressions at each level. Using expressions of themes instead of dynamisms at least partially alleviates the concern that Ss, especially those at the lower levels, will not give enough responses to make any statistical analysis possible. Also, the continuity of themes across levels makes them easier to rate than dynamisms. The fact that a satisfactory level of interrater reliability was obtained suggests that the expressions were adequately operationalized.

The data from this study offer empirical information bearing on the
theoretical conception of how the levels differ from one another. This information is observable in the averages for given distances and in the column averages on the bottom of Table I. Generally, the smaller the distance between levels, the more positive the correlations. The column averages show positive correlations between I and I-II, II and I-II, and even I and II. However, between II and II-III and II and III negative correlations exist. Then, between II-III and III the correlation is positive again. The theory, in fact, hypothesizes that the transitions are not all analogous. Specifically, the change from Level II to III is neither the same, nor “more of” the change from Level I to II. Piechowski (14, p. 265) explained the difference:

One can think of integration (Level I) and disintegration (Level II) as opposite poles of a continuum between maximum of structure and total lack of structure.

... (However,) the structure of unilevel disintegration (Level II) and the structure of multilevel disintegration (Level III, and above) are entirely different. In unilevel disintegration conflicts are horizontal, the opposing tendencies of equal value, everything is relative, arbitrary, governed by moment and circumstance. In multilevel disintegration, the conflicts are vertical, the opposing tendencies of lower and higher value... relativism and chance yield to a developmental hierarchy of autonomous direction and autonomous choice.

This, then, may explain why I and II are more similar than II and III. However, it will be recalled that the pattern of correlations obtained also indicates that in some respects I and III appear more similar than II and III. The levels do not form a continuum. Even though they are developmentally more distant, the theory does hypothesize certain structural similarities between Levels I and III: neither I nor III share Level II’s high degree of diffusion, lack of direction and need for belonging and approval. A similar problem occurred in Kohlberg’s research, where for a number of years stages 2 and 5 were often confused. The confusion ended with the introduction of stage 4½ (11).

The purposes of the present research was to begin the development and validation process for measures relevant to the theory. Because most of the measures achieved a good degree of success, some brief observations will be made regarding the methods themselves.

The nondirective instructions of the Autobiography method make it the least biasing method, in that it does not solicit responses related to themes. Nonetheless, each S responded on almost all the themes. The method did attain a fair measure of convergent validity which is surprising in light of the fact that it is rarely considered a valid assessment device for research. This study demonstrates that it may be the constructs and their theoretical and empirical significance which determine the soundness of the autobiographical approach rather than the method per se. The principal drawbacks of the A method are the time required to write the autobiography, the time and experience required to rate it, and that it will most likely be completed only by highly motivated Ss.

In contrast to rating the autobiographies, the rating of the definition-response (D) and the reasons (R) is considerably less taxing because these methods are free of unnecessary life history, the total number of responses is determined by the S and not by the rater, and they involve rating only the levels and not the themes; consequently, rating reliabilities tend to be higher. The D method, being more directive than the A method, is possibly more susceptible to eliciting biased responses. The heterolevel interrelationships were less regular with this method than with the others which may indicate that one or more of the theme descriptions was influencing Ss toward certain expressions. Nevertheless, Lysy (13) used the D method to assess level of development in research comparing personal growth of counselors and non-counselors. In her study of 42 Ss, she found very strong correlations between the level of development and various measures of developmental potential and its components.

One obvious advantage of the Situation-choice method is that the responses are predetermined and prerated, and therefore mechanically scorable. Another is that it demands less time from the S and therefore is a more flexible research tool. The drawbacks of this device are that it does not include choices representing the intermediate levels, and it is possibly more susceptible to the vagaries of social desirability than the A, D, and R methods.

Regarding the questionnaire (Q) method there are a number of possible reasons for why it failed to satisfy convergent validity requirements. One explanation is that it is much more open to the influence of social desirability than the other measures. The theory itself predicts that persons who respond to their world from a Level II perspective will be more susceptible to the influence of others’ opinions. Thus, Ss in different levels of development might be differentially affected by this contaminating factor. The individual S’s data are highly consistent with this explanation in that Ss who were lowest on the other measures appeared more likely to respond to the most socially acceptable (Level III) items on the questionnaire.
V. CONCLUSION

The present study accomplished its goal of developing more flexible assessment devices as alternatives to the autobiographical method. Three of the four new methods satisfied the convergent and discriminant criteria for establishing construct validity. By introducing the concepts of themes and their expressions at different levels, the study also succeeded in demonstrating a way to facilitate assessment of the levels of the theory of positive disintegration, thereby making the theory more amenable to future research.

APPENDIX A

Theme Definitions

Susceptibility to the Influence of Others entails being strongly affected by what others say and think. The effect of others' actions and opinions can be experienced in changes of behavior (to please, to attract attention, to conform) or in change of feelings (e.g., feeling liked or disliked). Dimensions which help to discriminate expressions of susceptibility at different levels are (a) whether the expression occurs automatically and without reflection or is self-chosen, and (b) to what end the influence serves (is it self-serving, conforming, rebellious, or self-critical?).

Personal Conflict refers to a real or imagined experience of a seemingly unsolvable personal dilemma regarding priorities, obligations, goals, personal aims, or moral values, which engenders feelings of tension, anxiety, desperation, or helplessness. The response to the conflict is usually a good indicator of the level of the theme (e.g., expedient resolution simply to relieve anxiety, incapacity for resolution (like feeling trapped with no exit), knowing what is morally right but not doing it).

Inferiority arises from a comparison between some aspect of one's perceived self and the analogous aspect of one's ideal self. The perceived self is evaluated to be inadequate in some critical way. The specific aspect compared (one's material possessions, social skills, personal development, etc.) and the way the response is expressed help determine the level of expression.

Dissatisfaction is an emotional reaction toward external or internal attributes of one's existence. Examples: a person may be annoyed at owning too little or not being attractive, he or she may be frustrated with having made a poor choice of career, or he may experience strong anger toward himself for failing to live up to his ideals.

Self-observation involves taking note of one's own behavior in a situation, being aware of emotional responses which may typically go unnoticed, or dissociating at least for a moment from one's physical self, one's thoughts, or one's feelings. Expressions of this theme vary from an observation concerning one's power over others, to becoming aware of one's ambivalence in a situation, to a self-critical evaluation of one's moral and personal responsibilities.

Personality Ideal involves a thought or image of oneself (not as currently perceived) that is wishfully fantasized or projected as an ideal toward which one strives. The particular qualities of the ideal self allow one to differentiate among the levels of expression of this theme.
APPENDIX B

DESCRIPTION OF THEMES IN DEFINITION-RESPONSE METHOD

Susceptibility to the Influence of Others: Think of times when you have been strongly affected by what others think of you or when you have compared yourself in some way to others.

Personal Conflict: Think of those questions which cause strong doubts within you, that frustrate you, and perhaps result in anxiety or depression. The problems should be limited to struggles which are internal (for example, philosophical, sexual, emotional), not struggles which are primarily external (for example, a purely economic problem).

Inferiority: Recall times when you have felt inadequate, unworthy, not good enough. Possibly you felt frustrated with what may have been lacking in yourself (abilities, skills, talents, personal qualities, etc.).

Dissatisfaction: Consider those situations which have caused you to feel frustration or anger toward yourself. They may have been over something you did and later regretted, as well as over something you feel you should have done, but did not do. Likewise, you could have become angered with yourself for having felt a certain way, or believing something you no longer feel is true.

Self-observation: Think if there have been any times when you have tried to stand back and look at yourself objectively. Upon what specific things did you reflect, if you did so?

Personality Ideal: Think of your “ideal self” and those qualities which you think are best for an ideal life. What attributes have you most dreamed of having?

APPENDIX C

SITUATIONS AND CHOICES FOR THE THEME

Susceptibility to the Influence of Others

Suppose for a moment that you have belonged to a large social organization for a number of years. They are now beginning the process of electing a new president. You have a good number of friends who are encouraging you to run because they feel you can do the best job. The only other likely candidate is a friend of yours, Janet, who has been waiting and planning to run for a long time. Certain aspects of the job are exciting to you: for example, participating in conferences around the country, meeting interesting people, and having the chance to do something for others. The drawbacks are that it would be time-consuming, there is no pay, and it is easy to make enemies as president.

While imagining yourself in this situation, answer whether or not you would decide to run for president. Briefly explain your first two choices.

A. No, the personal rewards wouldn’t really seem to offset the fact that it takes a lot of time and there is no pay.
B. Yes, I’d feel a sense of obligation toward those friends who really wanted me to run.
C. Yes, the excitement of the job would compensate for the drawbacks.
D. No, Janet obviously has her hopes set on the job and I would feel terrible if I beat her.
E. Yes, if my friends really think I should run, then I’d be silly not to.
F. No, I wouldn’t want to run because it takes too much time and energy.
G. No, I wouldn’t want to lose Janet’s friendship.
H. Yes, I feel Janet would realize I don’t think any less of her by my competing with her.
I. No, while I would want to confirm my friends’ feelings for me, I’d realize that it’s not what I really want to do.
J. Yes, I am confident that if I were elected, Janet and I could work something out so that she would feel included.
K. No, it would be intelligent of me to avoid the inevitable unpopularity that befalls all leaders.
L. Yes, being president would be an excellent way to show I can handle people.

REFERENCES