1975

Development of accuracy in self-perception

Brenda K. Holland

The University of Montana

Let us know how access to this document benefits you.

Follow this and additional works at: https://scholarworks.umt.edu/etd

Recommended Citation
https://scholarworks.umt.edu/etd/5754

This Thesis is brought to you for free and open access by the Graduate School at ScholarWorks at University of Montana. It has been accepted for inclusion in Graduate Student Theses, Dissertations, & Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.
THE DEVELOPMENT OF ACCURACY IN SELF-PERCEPTION

by

Brenda Griffith Holland

B.A., Lawrence University, 1970

Presented in partial fulfillment
of the requirements for the degree of
Master of Arts
University of Montana
1975

Approved by:

[Signatures]

Chairman, Board of Examiners

Dean, Graduate School

Date

July 31, 1975
The present study tested the hypothesis that self-perception accuracy increases with age. Five secondary hypotheses relating other characteristics to self-perception accuracy were also investigated. Participating in the study were 43 boys distributed into five groups representing grades 4, 6, 8, 10, and 12. Groups were similar in mean IQ and self-esteem. The investigation focused on three classroom behaviors - class participation, interpersonal talking, and class grades. These were defined behaviorally for the experimenter and participants. From the records from ten days of controlled classroom observation by the experimenter, subjects within each group were ranked on frequency of participation and frequency of talking. On the basis of teacher's final quarter grades, subjects within each group were ranked on grades. These experimentally determined ranks were the criterion against which the subjects' self-estimates were measured. Subjects then assigned themselves a rank within their peer group by filling out a comparison ranking questionnaire. This form defined the behavior and listed all the members of the group. Subjects indicated next to each name whether the person named or they themselves displayed the behavior more. Self-perception accuracy was defined as the absolute value of the difference between experimenter's ranking and subject's ranking.

One-way analysis of variance showed no significant differences between age groups. However, subjects showed a trend toward increasing accuracy with increasing age on participation and talking. No significant relationships resulted from the investigation of the secondary hypotheses. The tendency to over-rate oneself was unrelated to age. Neither IQ nor self-esteem was related to accuracy of self-perception. Peer-group rankings did not agree more with experimenter rankings as age increased. Self-perception accuracy was not consistent over the three traits investigated.

Because the test did not have an adequate ceiling, the study did not provide an adequate test of the hypotheses under consideration. Subjects showed an unexpectedly high degree of accuracy using a comparison-ranking self report. Even though the hypotheses were not supported, the study demonstrated the superiority of a little-used method of quantifying self reports - comparison ranking.
<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vi</td>
</tr>
</tbody>
</table>

Chapter

I. INTRODUCTION ........................................... 1

Present Study ........................................... 13

II. METHOD .................................................. 17

Standardized Self-Esteem Scale ........ 19
Self-Report Scale ......................... 22
Peer-Ranking Scale ....................... 22
Observer's Reports ..................... 24
Subjects ........................................... 26
Procedure ........................................ 28
Analysis .......................................... 29

III. RESULTS .............................................. 33

Main Hypothesis ............................... 33
Secondary Hypotheses ....................... 35

IV. DISCUSSION .......................................... 43

Accuracy of Comparison Ranking .......... 43
Secondary Findings ......................... 46
Implications for Further Study ........... 48

V. SUMMARY .............................................. 50

REFERENCES ........................................... 52
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.</td>
<td>Reliability of the Piers-Harris Scale</td>
<td>20</td>
</tr>
<tr>
<td>II.</td>
<td>Reliabilities of the Self-Report Scale</td>
<td>23</td>
</tr>
<tr>
<td>III.</td>
<td>Reliabilities of the Peer-Ranking Scale</td>
<td>25</td>
</tr>
<tr>
<td>IV.</td>
<td>Characteristics of the Experimental Groups</td>
<td>34</td>
</tr>
<tr>
<td>V.</td>
<td>Analysis of Variance of Accuracy of Five Age Groups</td>
<td>36</td>
</tr>
<tr>
<td>VI.</td>
<td>Tau Rank-Order Correlation Between Peer-Rankings and Observer's Rankings</td>
<td>41</td>
</tr>
<tr>
<td>VII.</td>
<td>Tau Rank-Order Correlation Between Teacher-Rankings and Observer's Rankings</td>
<td>41</td>
</tr>
<tr>
<td>VIII.</td>
<td>Frequency of Accuracy Scores</td>
<td>45</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figures</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Average Accuracy by Grade Level and Characteristic</td>
<td>37</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

The present study explored the relationship between age and accuracy of self-perception in boys from 9 to 17 years of age. Subjects ranked themselves in their own peer group on three behaviors - class participation, interpersonal talking, and class grades. To determine their accuracy, the first two self-rankings were compared with rankings arrived at through controlled classroom observation by the experimenter. To determine accuracy of self-ranking on grades, the self-rankings were compared with final grades for the quarter. The average accuracies of the different age groups were compared.

In recent years a great deal of research has been done exploring various aspects of the self concept. The self-concept can generally be defined to include all of an individual's perceptions and feelings about himself and his relationship to his environment. Only a small number of these recent self-concept studies concentrate on identifying stages or critical periods in the development of self concept, hence knowledge in the area of development of self-concept is at present sketchy. Evidence suggests that the discrepancy between ideal and real self and between social and real self increases from fifth to eighth
to eleventh graders as measured by a questionnaire administered with several different instructions (Katz and Zigler, 1967). Subjects filled out the questionnaire three times. They indicated how they really were, how they wished they were, and how they thought others saw them. From the fourth to seventh grade at least, a slight drop in the positiveness of self-evaluation on measures of personality traits may be contributing to the increased discrepancy (Amatora, 1957). Description of the ideal self has been shown to change from parental figures in early school years to glamorous heroes in middle years around fourth through sixth grades to a composite of desirable characteristics beginning in high school (Havighurst, 1946; Havighurst and MacDonald, 1956). Significant increase in self-concept stability from the fourth to sixth grade was found by Perkins (1958) in his comparison of subjects' Q-sorts over a two month period. Contrary to the evidence on continuing development of the self concept through high school years presented by several of the preceding studies, a two year longitudinal Q-sort study concluded that self-concept stability does not increase after the eighth grade (Engel, 1958).

An extensive study by Coopersmith (1967) approached the development of self concept by examining the child-rearing antecedents of high self-esteem children. This piece of research, using fifth graders, indicated that parental acceptance, clearly defined limits, respect for
individuality within the limits, and high parental self-esteem correlate significantly with high self-esteem in children of this age.

The largest number of published developmental studies concerning self concept deal with the development of accuracy in self-perception. Generally in these studies, an individual's self-perception is defined as consisting of all the descriptive statements that he believes can be correctly applied to himself. The environment continually provides information in such forms as successes, failures, and the labelings and reactions of others. The rate at which a child learns to assimilate this feedback into a relatively accurate body of self-perceptions cannot be described as yet with any certainty. Methodological problems with the studies on this question make interpretation of them difficult. Ruth Wylie (1961), in her book surveying self-concept research, devoted an entire chapter to review of studies on self-perception accuracy. She dealt both with self-perception studies which ask the subject to report how he sees himself and with social-perception studies which ask how he thinks that others see him. In either case, serious interpretive problems arise if several precautions are not taken. To begin with, the subject's report and the observer's report which is to be used as an accuracy criterion must answer the same question or deal with the same dimension. Where this
requirement is violated, lack of correspondence between subject's report and observer's report may simply result from the subjects' different understanding of the question asked and thus reflect no inaccuracy in self-perception.

Precautions must also be taken to insure that the subject and observer are using the rating scale or other response choices in the same way. If a 7 is average in the subject's use of the scale while a 5 is average for the observer, discrepancies between the two reports will not necessarily reflect the subject's inaccuracy. This problem is multiplied in studies where the subject must report how he thinks others see him. In this case he not only must use the rating scale exactly as the observers do, but he also must know how the others' opinions of him will average out in order to arrive at an accurate estimate.

Wylie further stressed that researchers should guard against the influence of stereotyped accuracy. If the instruments are so worded that subjects' or observers' evaluations would be about the same regardless of who is under consideration, then accuracy will result merely because of stereotyped responding. The questions must be formulated to allow adequate differentiation among people.

Several potentially relevant variables have gone uncontrolled in many studies of development in self-perception accuracy even though they have been frequently linked to self-perception. Numerous studies have related IQ to aspects
of the self concept including self-perception accuracy (Arseven, 1942; Brandt, 1958; Katz and Zigler, 1967).

Most notably, Holt's (1951) study using college students found a .77 correlation between IQ and self-perception accuracy.

Although the existence of a relationship between self-esteem, meaning the positiveness of one's self concept, and age is yet disputed, enough studies have found significant variations in self-esteem with age to warrant a control for self-esteem in any developmental study of accuracy of self-perception (Lin, 1963; Piers and Harris, 1964). Wylie criticized studies of accuracy in self-perception to date for failure to control for this variable since it has been demonstrated that self-perception accuracy is positively related to high self-esteem (Brandt, 1958).

Most studies published to date are flawed by several of these methodological errors pinpointed by Wylie. Bailey and Gibby (1971) studied the accuracy of self-perception as related to intelligence using 112 sixth graders and 124 twelfth graders. Each student was asked to rate his own intelligence on the Gibby Intelligence Rating Scale. The journal article fails to describe this scale except to mention that choices run from "retarded" through "genius." The subjects also indicated how they thought father, mother, teacher, friend and wishful thinking would rate them. The twelfth grade students were then given the Otis Quick-
Scoring Mental Ability Test because they did not already have one on file. The subject's own estimate of intelligence was converted to an IQ score and compared with his Otis IQ score. Results showed that twelfth graders' estimates of intelligence deviated from Otis IQ scores less (-5.99) than sixth graders' estimates (-8.35). Also, sixth graders showed significantly more variation in how they guessed others would rate them.

In Bailey and Gibby's study the subjects may well have been answering a different question than was asked by the Otis, which functioned as observer's report in this case. Intelligence as defined by the Otis and by the word "intelligence" to a school child may be different in a number of ways. Twelfth graders may well have used a definition closer to that of a standardized IQ test than did the sixth graders. Furthermore since a conversion system had to be devised to convert the subjects' ratings into Otis-type scores, there can be no assurance that the subjects' use of the scale was the same as the Otis meaning for it. The authors do report that 77% of the sixth graders and 82% of the twelfth graders rated their intelligence as "average." Such stereotyped responding often leads to an inflated accuracy correlation. A .46 correlation between twelfth grade self-ratings and the Otis score is considerably higher than other self-perception accuracies involving standardized IQ tests reported in the literature.
Previously published correlations range from .11 to .21 (Webb, 1955; Wolff, 1969). The two age groups had comparable mean IQ, but the similarity of their mean self-esteem scores was not established.

O'Hara and Tiedman (1959) also used a standardized test as the observer's report or accuracy criterion. Interested in vocational choice, they administered various standardized tests on work values, aptitudes, interests, social classes and general values to 1021 high school age boys. Then they asked each subject to rate himself on all the dimensions measured by the tests using the test manual's own definitions of what the test measured. The exact self-rating scale is not described. They found consistent increases in self-rating accuracy from freshman to senior year. Aptitude correlations increased from .44 to .69, while interest correlations increased from .70 to .83. O'Hara and Tiedman attempted at least to insure that the subject and the observer were answering the same question. There is still some doubt whether one sentence about what the test measures can communicate exactly what the test encompasses. No controls were included for IQ or self-esteem differences between the groups although the large size of the groups suggests that mean differences may not have been significant. The most serious problem with interpreting this experiment results from the statistical artifact brought in by correlating two evaluative self-
reports. Whatever their content, two evaluative self-reports made by the same individual tend to correlate positively (Wiley, p.237). This tendency may account for the high degree of correlation found between self-rating and standardized test score. It is difficult to determine if this statistical artifact exaggerated the age group differences in self-perception accuracy on vocational questions.

Phillips (1963) studied the accuracy of self-perception in the younger child. He administered a ten-item modification of Amatora's Children's Personality Scale (Wylie, 1961) to 96 third graders and 96 sixth graders. Responses were made on a three-point rating scale. The classroom teacher also rated each child with the scale as did three of each child's peers. Also included in the experiment was a simple level-of-aspiration task. After the first trial, the subjects were asked to show on a five-point scale how well they expected to do on the second. Phillips found a significant increase between third and sixth grade in the correlations between self-ratings and teacher ratings (.17 to .57) and also between self-ratings and peer ratings (.00 to .40).

One deficiency in Phillips's procedure is lack of control for IQ and self-esteem between the two age groups. Also some evidence has accumulated that teachers are not equally able to evaluate their students at different ages.
Ausubel, Schiff, and Gasser (1952), for instance, found a significant decline from third to seventh grade in teachers' ability to predict sociometric status. Since half the America scale's questions deal with social characteristics, this variation in teachers' ability to evaluate children of different ages may have affected Phillips's results. Peer estimates are subject to the same inadequacy. Ausubel, Schiff, and Gasser (1952) also discovered that the accuracy of a child's perception of others' status fluctuates significantly from age to age with a general trend toward increasing accuracy with increasing age. Thus peer estimates may not be a reliable accuracy criterion for a developmental study since they cannot be held constant over age.

On Phillips's level-of-aspiration task, significantly more third graders than sixth graders indicated that they would "do much better" the second time. Since no average difference in performance was found between trials, the author concluded that sixth graders were more accurate because they opted mainly for doing "a little better."

These conclusions are questionable. First a very important semantic question is left unanswered. Do third graders and sixth graders mean the same thing by "do much better?" Secondly the conclusions may not have been the same had the author compared individual prediction and performance rather than relying on class averages.
The Ausubel, Schiff, and Gasser (1952) study dealt with accuracy of social perceptions rather than with self-perception accuracy. Two classes of students from each of grades 3, 5, 7, 11, and 12 were asked to indicate how each other student felt about being their friend on a five-point scale. They also rated how much they liked each other student and how popular they thought each other student was on the same scale. Because of a marked tendency in all grades to use only the upper three points of the scale, the authors did not compute accuracies based on individual predictions. Rather they computed an average of how much the subject predicted others liked him and correlated it with the average of how much all other students did like him. Using this method of analysis they found no significant increases in sociometric accuracy between ages, but did get a gradual increase in accuracy correlation with age from .678 for third graders to a .893 for seniors. In the literature, accuracy correlations on personality traits generally are more in the range of .15 to .60 (Amatora, 1956; Renzaglia, 1962; Webb, 1952). The correlations resulting from the Ausubel, Schiff, and Gasser study were possibly high as a result of using averages from only a three-point scale. No matching was done between age groups for IQ or self-esteem. In fact the classes were not all from the same socioeconomic background. Analysis of raw scores by the experimenters indicated that the groups were most likely not samples from
De Jung and Gardner (1962) researched the same question as Ausubel, Schiff, and Gasser using slightly different methodology. They also used a five-point scale, but asked each child to define his own scale with references from his life. The person in the whole world he most liked, he wrote at the top of the scale and the least liked at the bottom. Using these reference points then he decided how much he thought he would turn to each of his classmates in time of trouble and also how he thought each other student was going to rate him. These same questions were put to two classes in each grade from fifth through twelfth. Using the average absolute discrepancy between how each subject was actually rated and how he predicted he would be rated, De Jung and Gardner found significant accuracy differences between the age groups using analysis of variance. A gradually decreasing discrepancy occurred from a difference between predicted and actual ratings of 2.18 for fifth graders to 1.16 for eleventh graders. Twelfth graders, however, showed a significant increase in discrepancy to 1.46.

De Jung and Gardner should have offered some evidence that the subjects defined and used their rating scales similarly. The possibility arises that the older children were more accurate simply because they defined their rating scales in a more uniform manner. The authors do not report
whether the full five-point range of the scale was used by the subjects. The groups were equated for mental age but not for self-esteem.

Richard Brandt's (1958) study of self-perception accuracy did not focus primarily on the developmental question, but his results give some information about this. To get the subjects' report, he asked them to rank themselves in their own peer group. They did this by predicting on a class list which peers would do better and which not as well as themselves on each task. The six tasks were arithmetic, spelling, vocabulary, broadjumping, baseball throwing, and strength of hand grip. Three classes of sixth graders and two of twelfth graders participated in the study. As an accuracy criterion, Brandt used actual subsequent performance on appropriate sections of the California Achievement Test Battery and the athletic tasks themselves. Brandt designed his experiment to find out whether between-individual differences in accuracy are greater than within-individual differences across several tasks. Using analysis of variance across all his subject groups, he did find significantly greater between-individual differences. Using exact accuracy scores, no significant difference in accuracy between the age groups was demonstrated. Using a measure of how accurately the subjects placed themselves in the correct quarter of the class, however, Brandt found significantly greater accuracy among the twelfth
graders.

By using standardized tests as an accuracy criterion on part of the tasks, a situation was created in which the subjects' and observer's reports may not have dealt with the same question. Since Brandt was not primarily interested in the developmental question, he did not match his groups on IQ, self-esteem or even familiarity. Thus the data his study provides on development of accuracy in self-perception are only suggestive for further research.

Present Study

The present study grew out of the need for more definite information on the development of accuracy in self-perception. Literature to date yields little data about the stages and trends that a child may pass through in the formation of a relatively accurate body of self-perceptions. Each study encompasses only a very limited age range. Due to differences in design and in the trait considered, no synthesis into a continuous developmental description is possible. Furthermore, methodological problems cloud the interpretation of most of the studies reviewed.

Summarizing the findings to date, Phillips (1963) discovered a significant increase from third to sixth grades in self-report accuracy on a diverse personality scale. Both Brandt (1958) and Bailey and Gibby (1971) found significant increases in self-perception accuracy from sixth grade to the last years of high school. Bailey
and Gibby concentrated on intelligence estimates, while Brandt averaged accuracy over a variety of academic and athletic skills. O'Hara and Tiedman (1959) covered all the high school years showing significant increases in self-perception accuracy in the vocational realm through these years. Thus information on the development of accuracy in self-perception is far from complete because studies cover the development years only very sketchily. Studies of social perception development are more inclusive in age range, but do not deal directly with the self-perception question.

The present study explored the development of accuracy in self-perception over a broader age range, avoiding the methodological problems that make interpretation of the literature to date difficult. None of the studies so far has matched age groups on self-esteem and only a few have equated average IQ. The present study used groups matched on both these characteristics. Half of the studies in the literature compute accuracy using subjects' reports and observers' reports that may not have been dealing with exactly the same question. The present study asked for reports on easily observed, clearly defined classroom behaviors so that both the subject and the observer could be evaluating the same experiences. Descriptions of the behaviors appear in Appendix A. Most of the previous studies have also failed to insure that the subjects and observers
were in fact using the rating scales in a similar manner. The present study used a peer-ranking procedure similar to that used by Brandt to avoid the difficulty of defining an abstract scale.

The subjects compared themselves with each classmate on the behaviors under consideration and thereby assigned themselves a rank in their peer group. Since the peer group served as a common reference scale for both the subjects and the observer, a high degree of similarity in the scale's use was achieved. Webb (1956) compared self-report reliabilities of such a comparison ranking, a group ranking including self, and a five-point rating scale. Using an eleven week test-retest procedure, comparison ranking attained a reliability of .69 (N=160), group ranking .34 (N=68), and rating scale .19 (N=95) with naval officers. Such a comparison ranking scale also circumvents the possibilities of stereotyped responding and skewed response distributions such as have resulted from scales used in the past. Peer-group comparisons allow for adequate response differentiation and make a concentration of scores at one end of the scale unlikely.

The present study was designed to test the first and main hypothesis that accuracy of self-perception increases gradually during years 9 through 17. Such a trend is suggested by previous studies dealing with the question even though methodological problems cloud their interpretation.
The data collected also allowed tests of the following hypotheses:

2. The tendency to overrate oneself is negatively related to age.

3. Accuracy of self-perception is positively related to IQ.


5. Peer-group rankings agree more with rankings determined through controlled observation as age increases.

6. Self-perception accuracy is consistent over several different traits.
CHAPTER II

METHOD

The present study tested for a relationship between accuracy of self-perception and chronological age in boys. A longitudinal design would admittedly have been best for studying the question at hand, but, as is so often the case, the time involved in a longitudinal study made it unfeasible. Thus the present study was conducted with a cross-section of boys ranging in age from 9 to 17 years.

A person holds perceptions of himself on an almost infinite range of characteristics. In order to study self-perceptions, a limited number of aspects of self-perception must be singled out for study. This study attempted to deal with self-perceptions in the realm of common behaviors. Three frequently occurring classroom behaviors were selected on the basis of how clearly they could be observed and recorded. The study focused on class participation, interpersonal talking, and class grades.

The subject's self-perceptions on these three behaviors were determined by asking him to rank himself in his own peer group on the frequency of occurrence of each behavior. The subject compared himself with each other student and indicated whether he or the other student displayed the behavior more. In this way he unknowingly assigned him-
self a rank within his peer group. On each scale the behavior was described by way of examples at the top of the page. The situation was defined as that classroom for the whole quarter.

The accuracy criterion for class participation and interpersonal talking was an observer's report based on controlled classroom observation. To arrive at a ranking of the subjects in each group on these two characteristics, each student was observed individually for 40 one-minute periods and his participation and talking behaviors during that time recorded. The most appropriate accuracy criterion for the self-report on grades was the teacher's grades themselves. The teacher is continually rating the students on academic achievement and giving them feedback on their relative success using his own criterion. Thus the teacher's grades offered an accuracy criterion with which the subjects had had many opportunities to become directly familiar.

Each subject was evaluated for IQ level and degree of self-esteem before the study began in order to match the age groups on these characteristics. This information was used also to explore relationships between these variables and accuracy of self-perception.

A separate questionnaire was administered to establish peer-group rankings because of the importance of this accuracy criterion in the literature. Hopefully the results of this study will be more easily synthesized with studies to
date since the common ground of a similar accuracy criterion is available to serve as a point of comparison.

**Standardized Self-Esteem Scale**

The Piers-Harris Children's Self-Concept Scale (Piers and Harris, 1969) was selected as a self-esteem measure because it was designed and standardized for a very wide age range. It consists of 80 simply stated yes and no items, designed for ages 9 through 18 years. The younger students require a maximum of 20 minutes to complete the test, while older students usually take considerably less time. The items are worded half negatively and half positively to reduce the likelihood of an acquiescent response set. No consistent differences in response between ages or sexes have been found except for indications that the variability of response decreases with age.

Concerning the reliability of the scale, the manual for the Piers-Harris Self-Concept Scale gives the information from two studies shown in Table I (Piers and Harris, 1969, p. 4). Information on the validity of the Piers-Harris Self-Concept Scale comes from several different sources. Mayer (1965) administered both the Piers-Harris scale and Lipsett Children's Self-Concept Scale to 98 children from special education classes. He found a correlation of .68. Cox (1966) found a correlation of -.64 between the Piers-Harris scale and the personal problems


TABLE I

RELIABILITY OF THE PIERS-HARRIS SCALE

<table>
<thead>
<tr>
<th>Study</th>
<th>Grade</th>
<th>Sex</th>
<th>N</th>
<th>Index</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pennsylvania</td>
<td>3</td>
<td>Girls</td>
<td>56</td>
<td>Kuder-Richardson</td>
<td>.90</td>
</tr>
<tr>
<td>Public School (Piers and Harris, 1964; 95 items)</td>
<td>6</td>
<td>Boys</td>
<td>63</td>
<td>&quot;</td>
<td>.93</td>
</tr>
<tr>
<td>Oregon Public School (Wing, 1966; 80 items)</td>
<td>5</td>
<td>Both</td>
<td>244</td>
<td>&quot;</td>
<td>.77</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Girls</td>
<td>56</td>
<td>4 mo test-retest</td>
<td>.72</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>Boys</td>
<td>66</td>
<td>&quot;</td>
<td>.71</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>Both</td>
<td>60</td>
<td>&quot;</td>
<td>.72</td>
</tr>
</tbody>
</table>
checked on the SRA Junior Inventory for 97 children in grades 6 through 9.

Correspondence between the Piers-Harris scale and teacher or peer ratings has generally been low. However, there is no reason to assume a high degree of accuracy in informal estimation of another's feelings. Piers (1965) found correlations between the Piers-Harris Self-Concept Scale and teachers' ratings from .06 for 54 fourth grade boys to .41 for 57 fourth grade girls. Cox (1966) reported a correlation of .40 between the Piers-Harris scale and teachers' ratings of superego strength. Peer ratings match Piers-Harris self-esteem scores slightly better, perhaps because they represent averages rather than single estimates. Piers found correlations with peer ratings ranging from .26 for 54 fourth graders to .49 for 58 sixth graders. In the Cox study, superego strength as estimated by peers correlated .42 with Piers-Harris Self-Concept Scale scores.

Two studies have shown institutionalized mental retardates to score significantly lower on the Piers-Harris scale than normals (Piers and Harris, 1964; Gorlow, Butler, and Guthrie, 1963). The authors of the scale present this information as a demonstration of the power of the Piers-Harris scale to discriminate groups commonly viewed as having different levels of self-esteem. Other studies have shown that non-institutionalized mental retardates do not
score significantly lower than normals on the scale.

Self-Report Scale

The format which was used in determining how each student viewed himself in relation to his peer group on the three behaviors under consideration appears in Appendix A. At the top of each questionnaire is a description of the behavior illustrated with examples. Under this are questions asking the subject to compare himself in turn with each other like-sexed classmate on the frequency with which they display the behavior. The subject must decide whether he or the classmate named displays the behavior under consideration more frequently. Standard instructions for administering the scale appear in Appendix F. The reliability of this questionnaire was assessed using a three week test-retest procedure. Reliabilities appear in Table II.

Peer Ranking Scale

The format for the peer ranking scale appears in Appendix B. A separate form was used to arrive at a peer group ranking of the students because on the self-report form all the judgments were made in relation to oneself. The peer ranking scale has the description of the behavior at the top. Under this is a list of like-sexed classmates. The subjects were asked to rank their classmates on the frequency with which they display the behavior under consid-
Table II  Reliabilities of the Self-Report Scale

<table>
<thead>
<tr>
<th>Behavior</th>
<th>Grade 4</th>
<th>Grade 5</th>
<th>Grade 7</th>
<th>Grade 8</th>
<th>Grade 12</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N Coefficient</td>
<td>N Coefficient</td>
<td>N Coefficient</td>
<td>N Coefficient</td>
<td>N Coefficient</td>
</tr>
<tr>
<td>Participation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys</td>
<td>29</td>
<td>.77</td>
<td>14</td>
<td>.96</td>
<td>14</td>
</tr>
<tr>
<td>Girls</td>
<td>12</td>
<td>.70</td>
<td>12</td>
<td>.98</td>
<td>7</td>
</tr>
<tr>
<td>Combined</td>
<td>41</td>
<td>.74</td>
<td>26</td>
<td>.97</td>
<td>21</td>
</tr>
<tr>
<td>Grades</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys</td>
<td>29</td>
<td>.89</td>
<td>14</td>
<td>.69</td>
<td>14</td>
</tr>
<tr>
<td>Girls</td>
<td>12</td>
<td>.61</td>
<td>12</td>
<td>.83</td>
<td>7</td>
</tr>
<tr>
<td>Combined</td>
<td>41</td>
<td>.79</td>
<td>26</td>
<td>.77</td>
<td>21</td>
</tr>
<tr>
<td>Talking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys</td>
<td>26</td>
<td>.86</td>
<td>11</td>
<td>.98</td>
<td>18</td>
</tr>
<tr>
<td>Girls</td>
<td>13</td>
<td>.89</td>
<td>13</td>
<td>.82</td>
<td>11</td>
</tr>
<tr>
<td>Combined</td>
<td>39</td>
<td>.87</td>
<td>24</td>
<td>.94</td>
<td>29</td>
</tr>
</tbody>
</table>

Correlations were determined using Pearson Product Moment Correlation between test and retest scores. Fourth through eighth grade correlations were based on results from two different classes combined using Z transformation. Three weeks elapsed between test and retest.
eration excluding themselves from the list. Standard instructions for this scale appear in Appendix F. The reliability of the peer ranking scale was determined using a three week test-retest procedure. Reliabilities appear in Table III.

Observer's Reports

For class participation and interpersonal talking, group rankings arrived at through 8 to 10 days of classroom observation were used as the accuracy criterion. The experimenter observed each subject in turn for one-minute periods and recorded all instances of participation and talking behavior as defined by the scale. Classroom observations conducted in this way by the experimenter on class participation by fourth graders over a three week period attained a correlation of .62 between first half and second half of observation periods. Each of sixteen boys was observed for a total of only 10 one-minute periods.

The inter-rater reliability of this observation method was checked using two raters observing 12 boys over a five day period. On class participation, the ranks determined by the two raters correlated .91 using Tau Rank Order Correlation. On interpersonal talking, the Tau Rank Order Correlation between the two raters was .81.

In recording participation, one score was assigned for each question asked in front of the class during the observation period. One score was given for any answer to a teacher's question even if given out of turn. Another
Table III Reliabilities of the Peer Ranking Scale

<table>
<thead>
<tr>
<th>Behaviors</th>
<th>Grade 4</th>
<th>Grade 5</th>
<th>Grade 7</th>
<th>Grade 8</th>
<th>Grade 12</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Coefficient</td>
<td>N</td>
<td>Coefficient</td>
<td>N</td>
</tr>
<tr>
<td>Participation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys</td>
<td>29</td>
<td>.81</td>
<td>14</td>
<td>.93</td>
<td>13</td>
</tr>
<tr>
<td>Girls</td>
<td>13</td>
<td>.67</td>
<td>14</td>
<td>.77</td>
<td>7</td>
</tr>
<tr>
<td>Combined</td>
<td>42</td>
<td>.75</td>
<td>28</td>
<td>.87</td>
<td>20</td>
</tr>
<tr>
<td>Grades</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys</td>
<td>29</td>
<td>.74</td>
<td>14</td>
<td>.98</td>
<td>13</td>
</tr>
<tr>
<td>Girls</td>
<td>13</td>
<td>.76</td>
<td>14</td>
<td>.69</td>
<td>7</td>
</tr>
<tr>
<td>Combined</td>
<td>42</td>
<td>.75</td>
<td>28</td>
<td>.92</td>
<td>20</td>
</tr>
<tr>
<td>Talking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys</td>
<td>28</td>
<td>.83</td>
<td>14</td>
<td>.54</td>
<td>34</td>
</tr>
<tr>
<td>Girls</td>
<td>15</td>
<td>.85</td>
<td>18</td>
<td>.56</td>
<td>15</td>
</tr>
<tr>
<td>Combined</td>
<td>43</td>
<td>.84</td>
<td>32</td>
<td>.55</td>
<td>49</td>
</tr>
</tbody>
</table>

Correlations were determined using Pearson Product Moment Correlation between test and retest scores. Fourth through eighth grade correlations were based on results from two different classes combined using Z transformation. Three weeks elapsed between test and retest.
score was given if the subject raised his hand or volunteered any kind of help to the teacher without raising his hand. If the hand remained in the air over 15 seconds, an additional score was given for each 15 seconds it was kept in the air. A total class participation score was determined by summing all participation scores over all observation periods.

In recording interpersonal talking, one score was given for each separate comment to a peer during the observation period. Should the comment run over 15 seconds, an additional score was given for each 15 seconds it lasted. Two scores were given each time the subject was reprimanded by the teacher by name for talking any time that the rater was making observations. A total interpersonal talking score was determined by summing all talking scores over all observation periods.

For class grades, the actual final grades for the quarter in the class were used as the accuracy criterion. Toward the end of the quarter, the students had received enough feedback from the teacher on their performance to understand the teacher's grading procedures.

**Subjects**

The study was conducted in the school system of Granby, Colorado, which serves approximately 625 families from several small nearby towns. The enrollment in the entire sys-
tem was 875 students with at least two classrooms of students at each grade level. The background of the students could be characterized as lower middle class with less than 3% of the population on public assistance and less than 3% in very high income brackets. In recent years the school system had experienced about an 18% turnover in its student body each year because of construction in the area.

Data were collected from boys in the fourth, sixth, eighth, tenth, and twelfth grades. One classroom containing from 10 to 15 boys was selected for participation at each of the grade levels. The classes were matched as closely as possible on IQ and self-esteem because these two variables have been shown to have enough influence on accuracy of self-perception to confound the relationship between accuracy and age. Those students who did not have a recent Lorge-Thorndike Intelligence Test on file were administered the test. Degree of self-esteem of all students in the grades under study was assessed with a standardized test before the experiment began. Only one twelfth grade class was available for the study so classes were selected from the other grades to match as closely as possible the average IQ and average self-esteem of the twelfth grade boys. Any subject testing in the extreme 10% in IQ or self-esteem was automatically discarded. To determine if the remaining groups were likely to be samples from the same population, T tests of the difference between means were
run between the highest and lowest groups on both IQ and self-esteem. One sixth grade subject with a low IQ score was discarded in order to obtain groups that tested to be samples from the same population with 95% confidence.

Procedure

The experiment required approximately 45 minutes of each class's time to complete the self-ranking and peer-ranking scales plus another 8 to 10 hours of classroom observation by the experimenter. Of necessity both boys and girls participated in the study, but complete data were collected only for the boys and only their test results were analysed. It was desirable to avoid asking the subjects to compare themselves with members of the opposite sex, since evidence suggests that the ability to compare oneself with members of the opposite sex varies with age (Ausubel, Schiff, and Gasser, 1952). Thus subjects were asked to compare themselves only with members of their own sex to avoid the complication of different cross-sex comparison ability with age. Only the boys were included in results to simplify the analysis of data, matching of groups on IQ and self-esteem, and observation periods.

First, all members of fourth, sixth, eighth, tenth, and twelfth grades were tested with the Piers-Harris Self-Concept Scale and the Lorge-Thorndike Intelligence Test if they did not already have it on file. One class was selected at each grade level for inclusion in the study on the basis
of average self-esteem and IQ. In the middle of the last quarter of the school year after enough time had elapsed to allow the students to become familiar with the other members of their class and with the teacher, the experimenter made classroom observations over ten consecutive days. The classroom teacher ranked the subjects on talking and participation on each of these ten days. Then the self-rating scale and the peer-ranking scale were immediately administered to each class in one sitting. The study and its results were discussed with each class as soon as results became available.

Analysis

The principle statistic used in the analysis of the data was one-way analysis of variance. The boys of each class were first ranked on each of the behaviors using the data obtained from the classroom observations and teachers' grades. Accuracy was defined as the absolute value of the difference between the subject's self-ranking and this criterion-measure ranking. Absolute value was used because direction of inaccuracy is not relevant to an answer of the main question asked. An analysis of variance for each of the three behaviors was run over the five age groups using this discrepancy score.

Different precautions and statistics were needed to test the five minor hypotheses. Tests for the relationship between accuracy of self-perception and self-esteem or IQ
could be run only if it was first established that no relationship existed between accuracy of self-perception and the ranking assigned by the observer's reports. Many studies (Brandt, 1958; Froehlich and Moser, 1954) have shown that there exists a tendency to overestimate oneself on socially desirable characteristics and underestimate oneself on undesirable ones. Assuming this to be the case, those subjects highest on the observer's rankings of socially desirable behaviors have a better chance of being accurate than those near the end of the list. Since those first on the observer's rankings are likely to test higher in self-esteem or higher in IQ in the case of grades, an accuracy advantage may by chance be going to those high in self-esteem or IQ. It could happen that those high in self-esteem or IQ are more accurate in self-perception than other subjects simply because they are first on the observer's rankings and have an unfair advantage toward accuracy. Thus any correlation between accuracy and self-esteem or between accuracy and IQ becomes difficult to interpret.

It was established by inspecting the data that no relationship existed between self-perception accuracy and placement on the observer's rankings. Those placed first on the observer's ranking did not have better accuracy scores than other subjects. Tests of all five of the minor hypotheses were then appropriate.

Pearson's Product Moment Correlation was used to test
four of these hypotheses using the following raw score formula:

$$r_{xy} = \frac{N(\bar{X}\bar{Y}) - \bar{X}\bar{Y}}{N(\bar{X}^2) - (\bar{X})^2 \cdot N(\bar{Y}^2) - (\bar{Y})^2}$$

To test the hypothesis that the tendency to over-rate oneself is negatively related to age, each subject's three accuracy scores were first summed. Because the direction of inaccuracy was important for this test, the positive sign for overestimation and negative sign for underestimation were left on the accuracy scores. Then a Pearson's Product Moment Correlation was computed between signed composite accuracy and chronological age.

To test the hypotheses that IQ and self-esteem are positively related to self-perception accuracy, each subject's three accuracy scores were again summed, but this time without sign. Direction of inaccuracy was of no importance here. Pearson's Product Moment Correlations were then run between composite accuracy and Lorge-Thorndike IQ scores and between composite accuracy and Piers-Harris self-esteem scores.

In considering the hypothesis that self-perception accuracy is consistent over several different traits, a series of correlations was computed. Accuracy on estimating interpersonal talking was correlated with accuracy in estimating grades. Accuracy in judging grades was correlated with accuracy in judging class participation. Accuracy of
self-ranking on class participation was correlated with accuracy of self-ranking on interpersonal talking.

The remaining hypothesis was tested using Tau Rank Order Correlation computed with the following formula:

\[ \gamma = \frac{4P}{n(n-1)} - 1 \]

where \( P \) stands for the number of numbers below each number on the list which is larger than the number under consideration and \( N \) stands for the number of ranks.

In order to determine if peer-group rankings agree more with rankings determined through controlled observation as age increases, one peer-ranking for each grade was determined by summing the estimates of all the subjects. The peer-ranking was then compared with the observer's ranking at each grade level using Tau Rank Order Correlation.
Main Hypothesis

In order to provide an adequate test of the main hypothesis that accuracy of self-perception increases with age, it was necessary to equate the different age groups on peer familiarity by testing in a very small school system. Results are therefore based on five small groups of boys selected to be statistically similar in IQ and self-esteem, but different in age. All classrooms contained either ten or eleven boys who all participated in the study. In order to achieve groups that were statistically similar, the results from some of the subjects were discarded. First eliminated were any subjects whose IQ score or self-esteem score was extreme enough to be in the upper or lower 10% of a normal population. One fourth grade subject was also discarded before beginning the analysis because his IQ score prevented the group average from being close enough to the group averages of the other groups. The groups finally selected to test the hypotheses displayed the characteristics shown in Table IV.

The first hypothesis that self-perception accuracy increases with age was tested using one-way analysis of variance. The accuracy score for each subject on each trait
Table IV  Characteristics of the Experimental Groups

<table>
<thead>
<tr>
<th>Grade</th>
<th>Number</th>
<th>Average Age</th>
<th>Average IQ*</th>
<th>SD IQ</th>
<th>Average Self-esteem**</th>
<th>SD Self-esteem</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>7</td>
<td>10.1</td>
<td>105.3</td>
<td>11.7</td>
<td>63.4</td>
<td>4.7</td>
</tr>
<tr>
<td>6</td>
<td>11</td>
<td>11.9</td>
<td>109.5</td>
<td>9.3</td>
<td>60.1</td>
<td>8.6</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>14.4</td>
<td>105.0</td>
<td>11.8</td>
<td>56.4</td>
<td>10.8</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>15.8</td>
<td>113.7</td>
<td>6.7</td>
<td>57.7</td>
<td>8.8</td>
</tr>
<tr>
<td>12</td>
<td>9</td>
<td>17.9</td>
<td>113.2</td>
<td>12.7</td>
<td>58.4</td>
<td>5.1</td>
</tr>
</tbody>
</table>

* Based on scores from Lorge-Thordike Intelligence Test
** Based on scores from Piers-Harris Self-Esteem Scale
was determined by taking the absolute value of the difference between the subject's self-ranking and the observer's ranking. The analysis of variance summarized in Table V showed no significant differences between age groups on any of the three characteristics considered. Trends toward increasing accuracy with age in judging participation and talking were clearly present as shown in Figure I, however, even though group differences did not reach significant levels.

On participation, the fourth and sixth grade subjects were less accurate than older students in estimating their own peer-group ranking. Fourth and sixth grade self-rankings correlated with observer's rankings only .55, while eighth, tenth, and twelfth grade rankings correlated with observer's rankings .72 using Pearson Product-Moment Correlation. On talking, the eighth grade group was more similar to the younger subjects in average self-perception accuracy. The fourth, sixth, and eighth graders' self-rankings correlated with the observer's rankings only .47, while tenth and twelfth graders' self-rankings correlated with observer's rankings .75 using Pearson Product-Moment Correlation. On grades, self-perception accuracy was independent of age with fourth graders ranking themselves about as accurately as did the twelfth graders.

Secondary Hypotheses

The data collected also allowed tests of five other
TABLE V

ANALYSIS OF VARIANCE
OF ACCURACY OF FIVE AGE GROUPS

<table>
<thead>
<tr>
<th>Behavior</th>
<th>Source of Variance</th>
<th>DF</th>
<th>Mean Square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Between</td>
<td>4</td>
<td>1.93</td>
<td>1.22</td>
</tr>
<tr>
<td>Participation</td>
<td>Within</td>
<td>38</td>
<td>1.53</td>
<td>1.22</td>
</tr>
<tr>
<td>Talking</td>
<td>Between</td>
<td>4</td>
<td>3.62</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td>Within</td>
<td>38</td>
<td>3.31</td>
<td>1.09</td>
</tr>
<tr>
<td>Grades</td>
<td>Between</td>
<td>4</td>
<td>4.14</td>
<td>2.04</td>
</tr>
<tr>
<td></td>
<td>Within</td>
<td>38</td>
<td>2.03</td>
<td>2.04</td>
</tr>
</tbody>
</table>

Significance at the .05 level requires $F=2.62$. 
Figure I

AVERAGE ACCURACY BY GRADE LEVEL AND CHARACTERISTIC

---

**Participation**

**Talking**

---

AVERAGE RANKS IN ERROR

---

Grades

---

GRADE LEVEL

---

GRADE LEVEL
hypotheses related to accuracy of self-perception. Inspection of the data showed that those subjects ranked at the extremes by the observer had slightly larger accuracy scores and thus were slightly less accurate than those ranked in the middle. Therefore subjects high in IQ or self-esteem showed no accuracy advantage that would bias tests on the remaining hypotheses.

The second hypothesis stated that the tendency to over-rate oneself is negatively related to age. It was predicted that older subjects would not over-rate themselves as much as younger subjects. Of the 129 judgments made by the subjects, 57 were underestimations, 56 were overestimations, and 16 were correct. Thus the data showed no propensity by the subjects to overestimate as is frequently reported in the literature. The accuracy score used to test this hypothesis was an average of the subject's accuracy scores on the three behaviors with a sign for direction. Pearson Product-Moment Correlation between age and accuracy was .18. Thus no negative relationship existed between age and overestimating in older subjects. Inspection of the data showed that the only group to overestimate more than underestimate themselves was the twelfth graders. In all 27 judgments made by this group, 7 were underestimations while 15 were overestimations. The data then give no support to the hypothesis that the tendency to over-rate oneself is negatively related to age.
Third to be considered is the hypothesis that accuracy of self-perception is positively related to IQ. It was predicted that those higher in IQ would be more accurate in their self-perceptions. The Pearson Product-Moment Correlation between IQ and accuracy averaged over the three behaviors was only -.08. Thus the data did not show any relationship between self-perception accuracy and intelligence as measured by standardized IQ tests.

The fourth hypothesis to be considered stated that accuracy of self-perception is positively related to self-esteem. It was predicted that those higher in self-esteem would be more accurate in their self-perceptions. The Pearson Product-Moment Correlation between self-esteem scores and accuracy averaged over the three behaviors was .38. This tendency for those high in self-esteem to be more inaccurate was not great enough to be significant at the .01 level. When plotted, the data fall into a trimodal curve with concentrations of highly inaccurate scores at low self-esteem (38), medium self-esteem (58), and high self-esteem (70). Although the correlation suggests a negative relationship between self-esteem and accuracy in self-perception, the relationship appears to be a complex one.

The fifth hypothesis stated that peer-group rankings agree more with rankings determined through controlled observation as age increases. Each subject assigned ranks to his peers on the peer-group ranking questionnaire. The
rankings made by all the subjects were added together to arrive at one peer-group ranking. Peer-group rankings were compared with the observer's rankings using Tau Rank Order Correlation. Although predominately positive the correlations are generally low and inconsistent as shown in Table VI. No age group did better than any other in matching observer's rankings and no behavior was consistently better estimated by the peer groups. Peer-group rankings did not become more accurate with age.

The teachers of the five classes were also asked to rank their students on participation and talking each day that the observer was present. Teacher-rankings of the groups were determined by summing the rankings from all the days. Tau Rank Order Correlation was again used to compare teacher-rankings with observer's rankings. Correlations appearing in Table VII ranged from .38 to .75 showing great differences between the teachers in their accuracy. Each teacher, however, maintained about the same level of accuracy from behavior to behavior.

Finally to be considered is the hypothesis that self-perception accuracy is consistent over several different traits. To test this hypothesis subjects' accuracy on each behavior was correlated with their accuracy on each other behavior using Pearson Product-Moment Correlation. Accuracy on participation correlated with accuracy on grades .05. Accuracy on participation correlated with accuracy on talking
### TABLE VI

**TAU RANK-ORDER CORRELATION**

**BETWEEN PEER-RANKINGS AND OBSERVER'S RANKINGS**

<table>
<thead>
<tr>
<th>Behaviors</th>
<th>4th</th>
<th>6th</th>
<th>8th</th>
<th>10th</th>
<th>12th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participation</td>
<td>.29</td>
<td>.49</td>
<td>.61</td>
<td>-.16</td>
<td>.56</td>
</tr>
<tr>
<td>Talking</td>
<td>.65</td>
<td>.57</td>
<td>.35</td>
<td>.79</td>
<td>.29</td>
</tr>
<tr>
<td>Grades</td>
<td>.61</td>
<td>.83</td>
<td>.56</td>
<td>.43</td>
<td>.56</td>
</tr>
</tbody>
</table>

### TABLE VII

**TAU RANK-ORDER CORRELATION**

**BETWEEN TEACHER-RANKINGS AND OBSERVER'S RANKINGS**

<table>
<thead>
<tr>
<th>Behaviors</th>
<th>4th</th>
<th>6th</th>
<th>8th</th>
<th>10th</th>
<th>12th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participation</td>
<td>.47</td>
<td>.75</td>
<td>.38*</td>
<td>.47</td>
<td>.70</td>
</tr>
<tr>
<td>Talking</td>
<td>.56</td>
<td>.68</td>
<td>.64*</td>
<td>.53</td>
<td>.65</td>
</tr>
</tbody>
</table>

*Incomplete data turned in by this teacher.*
.03. Accuracy on grades correlated with accuracy on talking
.16. None of these correlations show any significant rela-
tionship. Thus the data gave no support to the hypothesis
that an individual's self-perception accuracy is consistent
over different traits.
CHAPTER IV

DISCUSSION

Accuracy of Comparison Ranking

The method of self-estimation used in this study resulted in more reliable and more accurate self-judgments than reported by other studies using a standardized accuracy criterion such as standardized tests or controlled observation. Webb (1955) reported a correlation of .21 between self-estimates and the Otis IQ test. Torrance (1954) found self-estimates correlating only .22 to .41 with achievement tests and Arsenian (1942) got self-estimates correlating .30 to .57 with college entrance tests. In the present study subjects' self-rankings on participation correlated with observer's rankings .65, self-rankings on grades correlated with observer's rankings .62, and self-rankings on talking correlated with observer's rankings .47.

Very high accuracy correlations reported in the literature often can be traced to stereotyped responding. Subjects tend to use only the middle of the scale. Scores assigned to subjects on the accuracy criterion measure are also normally distributed and so concentrate in the middle of the scale. An artificially high correlation coefficient results. In the present study subjects did arrive at middle rankings more frequently as they individually compared themselves with each of their peers. However, the
observer ranked the subjects 1 through 10 so that ranks assigned fell evenly over the full scale. Thus the high accuracy displayed by the subjects could not be the result of stereotyped correspondence. The high degree of accuracy by the subjects is more likely to be the result of providing a very structured, clearly defined procedure for self-estimation. Most previous studies have used some type of abstract rating where subjects must rate themselves on a numbered scale. A subject participating in the present study simply had to decide whether he or a peer displayed some common clearly defined behavior more often.

Each self-ranking could have been from 5 to 9 ranks in error depending on the rank assigned to the subject by the observer. A subject assigned the rank of 10 could guess himself to be 1 and get a +9 accuracy score. A subject assigned the rank of 6 could be +5 ranks in error at most by guessing himself to be 1. Thus the degree of possible inaccuracy depended on the rank assigned by the observer. The actual accuracies of the subjects distributed themselves as appears in Table VIII. About half of the judgments were either correct or only one rank in error. This degree of accuracy is quite remarkable when considering the difficulty of many of the discriminations. Subjects were often very close to being the same in grades or in frequency of talking or in frequency of participating. These near ties made it very difficult for subjects to make correct choices all of the time. However, the results of this study show that child-
<table>
<thead>
<tr>
<th>Accuracy (Self-ranking minus Observer-ranking)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency of Each Accuracy Score</td>
<td>16</td>
<td>44</td>
<td>28</td>
<td>18</td>
<td>14</td>
<td>7</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>129</td>
</tr>
</tbody>
</table>
ren even as young as the fourth grade are capable of making highly accurate self-estimates provided that the judgment process is structured for them and the characteristics are concretely and behaviorally defined. When clear and frequent feedback was given as with grades, fourth graders were just as accurate as twelfth graders in judging their class rank. When the environment provided less definite information as in the case of class participation and interpersonal talking, then older subjects were more accurate. The data suggest that self-perception accuracy does increase with age especially as the environment provides less structured feedback. However, the high degree of accuracy by all subjects in this study made it impossible for age differences to emerge at a significant level.

Secondary Findings

Some useful information comes out of the investigation of the five secondary hypotheses of the study.

2. Although the literature generally reports more frequent overestimates of oneself and Brandt (1958) found substantially more overestimating using a procedure of self-evaluation similar to the one used in this study, the present study revealed overall equal incidence of overestimating and underestimating.

3. The present study found no relationship between IQ and self-perception accuracy. It is always difficult to get a high correlation coefficient if using subjects that are
similar on the characteristics under consideration. Eliminating subjects in the extreme 10% of the IQ range, therefore, made any relationship between IQ and self-perception accuracy more difficult to detect especially when the relationship may be very small as is suggested by other studies.

4. Again eliminating those subjects testing in the extreme 10% on self-esteem made any relationship between self-esteem and self-perception accuracy more difficult to substantiate using correlational techniques. The correlation of .38 between self-esteem and self-perception accuracy does not reach significance at the .01 level. The relationship between self-esteem and self-perception accuracy found here is not a linear one and so would require further investigation to be clearly interpreted.

5. The present study raises serious question about the use of peer-group evaluations or teacher evaluations as a standard against which to measure self-perception accuracy. Peer-group rankings were highly inconsistent and inaccurate. Some teacher-rankings were accurate while other were not. Teacher evaluations could only be of value as an accuracy measure if the individual teacher's ability to make accurate evaluations was first substantiated.

6. Although self-perception accuracy was not consistent over the behaviors of class participation, interpersonal talking, and class grades, this lack of consistency may have resulted from the restricted range of accuracy scores. Seventy per cent of the accuracy scores were two ranks in
error or less. With such a high degree of accuracy, the differences in accuracy scores from one self-estimate to another were too slight to be meaningful. The lack of correlation between accuracies on the different behaviors most likely reflects the chance nature of these small variations in accuracy. The results of the study do not provide an answer then to the question whether self-perception accuracy is consistent from trait to trait.

Implications For Further Study

Because all age groups participating in this study showed a very high degree of self-perception accuracy, this study did not adequately test the hypothesis that self-perception accuracy increases with age. Fourth graders were so accurate on the average in their perceptions that the twelfth graders could not do much better. In order to improve the design to provide an adequate test of the hypotheses, the range of the accuracy scores must be increased. This could be done by increasing the number of subjects in each experimental group. If each subject had to compare himself to 20 peers rather than to 10, he would be less likely to place himself within one or two ranks of the observer's placement. The self-ranking would be no more difficult to complete, but more chances for variation or error would exist with twice as many comparisons. If enough errors were made so that there was room for improvement, then differences between younger and older subjects could
emerge if they actually do exist.

Although the present study did not provide significant evidence that self-perception accuracy increases with age, it did demonstrate the usefulness of a relatively little-used procedure for obtaining self reports. Individual comparison with peers results in highly reliable evaluations. Comparison ranking also provides estimates that agree closely even in young children with externally controlled observations. Obtaining self reports in quantified form has presented great problems in self-concept research. Rating scales have been unsatisfactory because they are not understood and used uniformly. Often response choices are too narrow to express the subject's opinion with the result of stereotyped responding. Using the comparison ranking procedure, the subject can quantify his self report in terms that are meaningful to him. This type of self report is obviously useful for research on self-perception accuracy especially when used in conjunction with controlled observation for an observer's report. However, comparison ranking could easily be adapted to many kinds of self-concept studies with great benefit to research in this area.
CHAPTER V

SUMMARY

This study tested the hypothesis that self-perception accuracy increases with age. Five secondary hypotheses relating other characteristics to self-perception accuracy were also investigated.

Subjects were 43 boys distributed into five groups representing grades 4, 6, 8, 10, and 12. Groups were similar in mean IQ and self-esteem. The investigation focused on three classroom behaviors - class participation, interpersonal talking, and class grades. These were defined behaviorally for the experimenter and participants. From the records from ten days of controlled classroom observation by the experimenter, subjects within each group were ranked on frequency of participation and frequency of talking. On the basis of teacher's final quarter grades, subjects within each group were ranked on grades. Subjects then assigned themselves a rank within their peer group by filling out a comparison ranking questionnaire. This form defined the behavior and listed all the members of the group. Subjects indicated next to each name whether the person named or they themselves displayed the behavior more. Self-perception accuracy was defined as the absolute value of the difference between experimenter's ranking and subject's ranking.
Although subjects showed increasing accuracy with increasing age on participation and talking, one-way analysis of variance showed no significant differences between age groups. No significant relationships resulted from the investigation of the secondary hypotheses. The tendency to over-rate oneself was unrelated to age. Neither IQ nor self-esteem was related to accuracy of self-perception. Peer-group rankings did not agree more with experimenter rankings as age increased. Self-perception accuracy was not consistent over the three behaviors investigated.

Subjects showed an unexpectedly high degree of accuracy. Nearly half the judgments were correct or only one rank in error. Because the test did not have an adequate ceiling, the results of the study were not conclusive. However, the usefulness of the comparison ranking method of obtaining self-reports was demonstrated.
REFERENCES


Arsenian, S. Own estimate and objective measurement. *Journal of Educational Psychology,* 1942, 33, 291-302.


Beilin, H. The prediction of adjustment over a four year interval. *Journal of Clinical Psychology,* 1957, 13, 270-274.


APPENDIX A
SELF-RANKING QUESTIONNAIRE: PARTICIPATION

NAME ____________________________

WHO PARTICIPATES THE MOST IN THIS CLASS?

You will know because he will
- ASK QUESTIONS
- ANSWER THE TEACHER'S QUESTIONS
- RAISE HIS HAND AND VOLUNTEER TO HELP

1. ___ RICK or ___ ME
2. ___ KURT or ___ ME
3. ___ TIM or ___ ME
4. ___ DEAN or ___ ME
5. ___ BOB or ___ ME
6. ___ RONNIE or ___ ME
7. ___ JODY or ___ ME
8. ___ GARY or ___ ME
9. ___ MANUEL or ___ ME
10. ___ FRANK or ___ ME
11. ___ ROSS or ___ ME
12. ___ CURTIS or ___ ME
SELF-RANKING QUESTIONNAIRE: GRADES

NAME _______________________

WHO GETS THE BEST GRADES IN THIS CLASS?

You will know because he will
- Get the best grades on his tests
- Get the best grades on his papers
- Get the best grades on his projects

1. ___ RICK or ___ ME
2. ___ KURT or ___ ME
3. ___ TIM or ___ ME
4. ___ DEAN or ___ ME
5. ___ BOW or ___ ME
6. ___ RONNIE or ___ ME
7. ___ JODY or ___ ME
8. ___ GARY or ___ ME
9. ___ MANUEL or ___ ME
10. ___ FRANK or ___ ME
11. ___ ROSS or ___ ME
12. ___ CURTIS or ___ ME
SELF-RANKING QUESTIONNAIRE: TALKING

NAME ________________________

WHO TALKS THE MOST TO OTHER KIDS WHEN IN THIS CLASS?

You will know because he will
- TALK TO OTHERS ON THE WAY IN AND OUT OF CLASS
- TALK MORE THAN THE OTHERS WHEN IN A GROUP OF KIDS
- SOMETIMES GET IN TROUBLE FOR TALKING TOO MUCH

1. ___ RICK or ___ ME
2. ___ KURT or ___ ME
3. ___ TIM or ___ ME
4. ___ DEAN or ___ ME
5. ___ BOW or ___ ME
6. ___ RONNIE or ___ ME
7. ___ JODY or ___ ME
8. ___ GARY or ___ ME
9. ___ MANUEL or ___ ME
10. ___ FRANK or ___ ME
11. ___ ROSS or ___ ME
12. ___ CURTIS or ___ ME
APPENDIX B
WHO PARTICIPATES THE MOST IN THIS CLASS?

You will know because he will
- ASK QUESTIONS
- ANSWER THE TEACHER'S QUESTIONS
- RAISE HIS HAND AND VOLUNTEER TO HELP

Put a 1 by the person who participates the most.

___ RICK
___ KURT
___ TIM
___ DEAN
___ BOW
___ RONNIE
___ JODY
___ GARY
___ MANUEL
___ FRANK
___ ROSS
___ CURTIS
WHO GETS THE BEST GRADES IN THIS CLASS?

You will know because he will
- GET THE BEST GRADES ON HIS TESTS
- GET THE BEST GRADES ON HIS PAPERS
- GET THE BEST GRADES ON HIS PROJECTS

Put a 1 by the person who gets the very best grades.

___ RICK
___ KURT
___ TIM
___ DEAN
___ BOW
___ RONNIE
___ JODY
___ GARY
___ MANUEL
___ FRANK
___ ROSS
___ CURTIS
WHO TALKS THE MOST TO OTHER KIDS WHEN IN THIS CLASS?

You will know because he will
- TALK TO OTHERS ON THE WAY IN AND OUT OF CLASS
- TALK MORE THAN THE OTHERS WHEN IN A GROUP OF KIDS
- SOMETIMES GET IN TROUBLE FOR TALKING TOO MUCH

Put a 1 by the person who talks to other kids the most.

___ RICK
___ KURT
___ TIM
___ DEAN
___ BOW
___ RONNIE
___ JODY
___ GARY
___ MANUEL
___ FRANK
___ ROSS
___ CURTIS
APPENDIX C
<table>
<thead>
<tr>
<th>ASKS QUESTIONS</th>
<th>ANSWERS TEACHER'S QUESTIONS</th>
<th>RAISES HAND VOLUNTEERS HELP</th>
</tr>
</thead>
</table>

**INTERPERSONAL TALKING**

<table>
<thead>
<tr>
<th>ON WAY IN AND OUT</th>
<th>DURING CLASS</th>
<th>GETS IN TROUBLE FOR TALKING</th>
</tr>
</thead>
</table>

**NAME** ______________________
EXPLANATION OF THE PROJECT

"I'm Mrs. Holland. I'm working on a psychological experiment and this class is going to help me with part of it. As you all know, people have learned a lot of very valuable things in the last hundred years through scientific experiments. As a result we now have landed on the moon, invented several wonder drugs that make most sicknesses less serious, and improved communication to the point where TV can bring us instant pictures of things happening anywhere on the earth.

There is one thing that we haven't learned much about yet and that is people themselves. We actually know very little about why people do what they do and feel the way they do. If we knew more about people we might be able to help a lot of people make happier lives for themselves.

You are being asked to participate in a scientific experiment. The purpose of the experiment is to learn more about the way people of different ages feel about themselves and others. It is important that you be completely honest in all the tasks you will be asked to do. Some very valuable knowledge could possibly come out of the results of this study."
STANDARD INSTRUCTIONS FOR THE PIERS-HARRIS SCALE

"You are being asked to participate in a research project by completing the questionnaire which you now have in front of you. The purpose of the study is to find out more about how students of your age feel about themselves. This booklet contains only questions about how you feel about yourself. It is not a test and so there are no right or wrong answers. Open your booklet now to the first page and fill in your name, age, and the other spaces you see there. Also tell what grade in school you first went to school in Granby. Now we will read the instructions at the beginning of the first page together. You follow along while I read them. ... We are interested in finding out how you really feel about yourself. No one will see your answers but me, so be as honest as you can. Do not answer the questions according to what you think you ought to be, but try to tell how you really feel you are today. You may begin now. (Fourth grade - We will begin reading the questions together now. Mark your answers as we go along.) Take as much time as you need to show how you really feel about yourself. Close your book and wait for the others when you finish."
STANDARD INSTRUCTIONS FOR
THE SELF-RANKING AND PEER-RANKING SCALES

"You can start by writing your name at the top of the page in the space provided. Do this on all six pages of your booklet. Now find your own name in the list of names below and cross out that whole question on all six pages. You should all now have the question with your own name in it crossed out on all six pages. You do not have to answer that one. The question to think about on this page is "Who participates the most in this class?" You will know who that is because he will ask questions, answer the teacher's questions, raise his hand and volunteer to help and things like that. Look at question number 1. Here you are to decide whether you participate more or the other person named in question 1 participates more. If you decide that you participate more, put an "X" in front of "me." If you decide that the other person participates more put an "X" in front of his name. In question number 2 you do the same thing again. You decide whether you participate more or the other person named in question 2 participates more. You give one answer to each of the questions the same way. Are there any questions about how you are to do this page? I'm the only one who will see your paper so try to be honest. Don't look at anyone else's paper because I want to know what you think. One more thing before you start. Consider
each answer carefully and think especially about the last two weeks that I have been in here. Stop when you finish the first page and we'll wait for everyone to go on together.

Is there anyone who needs more time for page one? Let's all turn to page two. The question to think about here is "Who gets the best grades in this class?" You will know who that is because he will get the best grades on tests, papers, projects, or whatever your teacher grades in this class for the whole quarter. You do this page the same way you did the last one. Are there any questions? Stop when you finish page two.

Is there anyone who needs more time for page two? Let's all turn to page three. The question to think about here is "Who talks the most to other kids when in this class?" You will know who that is because he will talk to other kids on the way in and out of class, talk more than the others when in a group of kids, sometimes get in trouble for talking too much, and things like that. Think especially about the last two weeks that I have been in here. You do this page the same as the others.

Is there anyone who needs more time for page three? Let's all turn to page four. The questions on the last three pages of the booklet are the same, but you are to answer them in a different way. The question on this page is "Who participates most in this class?" You are to decide who participates the very most and put a number 1 in front
of his name. Put a two by the person who participates the second most and a three for the person who participates the third most and so on. The boys should have one through ___ because you crossed out your own name and the girls should have one through ___ because you crossed out your own name.

Are there any questions about what you are to do? Remember, put a 1 by the person who participates the very most. Go ahead and complete the whole booklet this way. (Fourth and sixth will proceed page by page in this booklet also.)