

CHAPTER II

SCALE CONSTRUCTION

Richard Christie

A pool of 71 items believed to be theoretically congruent with statements based on *The Prince* and *The Discourses* was inherited from the year at the Center. It was not clear whether or not responses to them would be scalable. However, the apparent fit of these items with implicit assumptions made by power theorists and their relevance to informal appraisals of power figures in the behavioral sciences was encouraging. The temptation to find out whether or not the items would hang together empirically was too great to resist.

The 71 items (Table II-1) were placed in a Likert format. Respondents' options for each item varied from strong disagreement through some disagreement, indifference, or inability to make up one's mind, some agreement, to strong agreement. When agreement meant acceptance of statements worded in the Machiavellian direction, the scoring was from one for strong disagreement to five for strong agreement. When the statement was phrased so that agreement presumably meant acceptance of a non-Machiavellian point of view, the scoring was reversed. Throughout the remaining chapters high scores reflect a respondent's acceptance of a Machiavellian viewpoint, and low scores a rejection of such an outlook. The items were presented in a questionnaire dubbed "Mach II" and respondents were asked to indicate the response alternative closest to their opinion for each item. (The original items, before editing, had been referred to as "Mach I.")

At this point, it seemed best to find out if the items discriminated in any available population. College students were the most accessible captive audiences. Given the difficulties of obtaining a representative sample, it appeared advisable to obtain a diverse sampling of college students. David Gold administered the scale to 518 students in sociology courses at the State University of Iowa; Robert Agger administered it to 427 students in political science courses at the University of North Carolina; Harold Yuker administered it to 251 students in psychology at Hofstra College on Long Island. Not all of these 1196 students were majors in these disciplines and ideally it would have been better to have included respondents who were not taking courses in the behavioral sciences. However, at that time there were no funds on hand to test a representative sample of college students or any other population so we were grateful for such regional and course variety as was available through the cooperation of friends.

TABLE II-1

| Identification | |
|----------------|---|
| Tactics + | 1. A white lie is often a good thing.* |
| Views + | 2. Generally speaking, most people do not truly believe in anything new until they have experienced it. |
| Tactics - | 3. If one is morally right, compromise is out of the question.* |
| Tactics + | 4. There is no point in keeping a promise if it is to your advantage to break it.* |
| Views - | 5. Any normal person will stand up for what he thinks is right even if it costs him his job.* |
| Morality - | 6. No one has a right to take his own life.* |
| Tactics + | 7. Every time one problem is avoided another arises.* |
| Views + | 8. Nature has so created men that they desire everything but are unable to attain it.* |
| Tactics - | 9. Once a way of handling problems has been worked out it is best to stick with it. |
| Views - | 10. Most men like to tackle new and difficult problems.* |
| Tactics + | 11. Anyone who completely trusts anyone else is asking for trouble.* |
| Morality - | 12. There is too much of an emphasis upon practical subjects in our schools and not enough attention is paid to the moral teachings of great historical figures.* |
| Views - | 13. If people could understand one another's position, there would be no good reason for human problems. |
| Tactics - | 14. It is best not to let people think that you like them when you really can't stand them. |
| Views + | 15. It is easy to persuade people but it is hard to keep them persuaded. |
| Tactics + | 16. If there is any chance that a recommendation might backfire, be very cautious in recommending anyone. |
| Views - | 17. You really can't tell much about a man by the company he keeps. |
| Tactics + | 18. It is a good policy to act as if you are doing the things you do because you have no other choice. |
| Views - | 19. Most people who get ahead in the world lead clean moral lives.* |
| Views + | 20. Some of the best people have some of the worst vices.* |
| Views + | 21. It is safest to assume that all people have a vicious streak and it will come out when they are given a chance.* |
| Tactics + | 22. It is foolish to take a big risk unless you are willing to go the limit.* |
| Views - | 23. The power of one man over another is rarely the result of trickery or force.* |
| Morality - | 24. The world would be in much better shape than it is if people acted upon basic ethical principles. |
| Tactics - | 25. One should take action only when sure it is morally right.* |
| Tactics + | 26. The best way to settle an argument is for people to forget their differences of opinion as to what is right and wrong and to compromise on the basis of facts.* |
| Tactics + | 27. Humility not only is of no service but is actually harmful.* |
| Views + | 28. Generally speaking, men won't work hard unless they're forced to do so.* |
| Tactics + | 29. It is wise to flatter important people.* |
| Views + | 30. It is hard to get ahead without cutting corners here and there.* |

TABLE II-1 (continued)

Identification

| | |
|------------|---|
| Morality + | 31. People suffering from incurable diseases should have the choice of being put painlessly to death.* |
| Views - | 32. Even the most hardened and vicious criminal has a spark of decency somewhere within him. |
| Tactics - | 33. Once a decision has been made, it is best to keep changing it as new circumstances arise. |
| Views + | 34. Most people don't know what is best for them.* |
| Views + | 35. Never trust anyone who has a grudge against you.* |
| Views + | 36. One who deceives will always find those who allow themselves to be deceived.* |
| Tactics - | 37. It is best to give others the impression that you can change your mind easily. |
| Morality + | 38. History teaches no moral lessons; evil and good have always been present.* |
| Views - | 39. Most men will fight back when insulted. |
| Tactics - | 40. Generally speaking, it is best to say what you think even if it bothers others. |
| Views + | 41. When you come right down to it, it's human nature never to do anything without an eye on one's own advantage.* |
| Tactics + | 42. It is better to compromise with existing evils than to go out on a limb in attacking them.* |
| Morality + | 43. A good government should keep the people peaceful and obedient. |
| Views - | 44. Most men are brave.* |
| Tactics + | 45. The best way to handle people is to tell them what they want to hear.* |
| Views + | 46. The biggest difference between most criminals and other people is that criminals are stupid enough to get caught.* |
| Tactics - | 47. Honesty is the best policy in all cases.* |
| Views - | 48. Men are quicker to praise than they are to blame.* |
| Morality - | 49. It is better that a hundred criminals go free than one innocent man should be unjustly punished. |
| Tactics + | 50. It is easier to take advantage of someone you love than someone you fear.* |
| Views - | 51. Barnum was very wrong when he said that there's a sucker born every minute.* |
| Tactics + | 52. It is safer to be feared than to be loved.* |
| Views - | 53. Most people are basically good and kind.* |
| Tactics + | 54. Friends should be chosen with an eye toward what they might be able to do for you.* |
| Morality + | 55. Deceit in the carrying on of war is praiseworthy and honorable.* |
| Views + | 56. The great majority of men are more satisfied with what seems true than with the truth.* |
| Tactics + | 57. If a friend asks for advice, it is smart to think about what will happen if your advice backfires.* |
| Views + | 58. Men resent an injury at the hands of others more than one that is self-inflicted. |
| Tactics - | 59. When you ask someone to do something for you, it is best to give the real reasons for wanting it rather than giving reasons which might carry more weight.* |
| Views - | 60. It is possible to be good in all respects.* |

TABLE II-1 (continued)

| Identification | |
|----------------|--|
| Tactics - | 61. It is not a good idea to bring pressure upon people if you want them to do something.* |
| Morality + | 62. The most important thing in life is winning. |
| Views + | 63. Most men forget more easily the death of their father than the loss of their property.* |
| Tactics + | 64. One should upset as few people as possible while making decisions. |
| Tactics + | 65. It is a good working policy to keep on good terms with everyone. |
| Tactics + | 66. Just about anything one does can be justified after it is done.* |
| Tactics + | 67. Never tell anyone the real reason you did something unless it is useful to do so.* |
| Tactics - | 68. There is no excuse for lying to someone else.* |
| Morality + | 69. The most important thing in history is who won and now how the winning came about. |
| Morality - | 70. All in all, it is better to be humble and honest than to be important and dishonest.* |
| Views + | 71. Most people are more concerned with making a good living than with satisfying their conscience.* |

*The 50 items which discriminated at the .05 level in all three school populations.

Scores were summed over the 71 items and the distribution of total scores was separately tabulated for the respondents from each school. Each distribution was divided at the median to define high and low scorers within each. All responses to each item were tabulated and summed for each school separately and each of the three distributions was cut as close to the median as possible. This permitted the construction of fourfold tables pitting individual items against summed scores. Phi coefficients were then computed to obtain part-whole correlations of each item with subjects' scores on the total scale in each aggregate.

Fifty of the 71 items discriminated between high and low scorers on the total scale at the .05 level in all three populations, 13 other items discriminated in two of the three groupings, and five discriminated in one of the three. Only three items (9, 33, 37) failed to discriminate in any of the samples. One of the latter (item 33) showed a negative correlation in all three groups but it was significantly so in only one sample. All three of these items have a common theme, consistency versus inconsistency of behavior. They were intended to be variations on Machiavelli's observation, "I certainly think that it is better to be impetuous than cautious, for fortune is a woman, and it is necessary, if you wish to master her, to conquer her by force...". In this case, it is difficult to determine whether our items were so ineptly worded that they missed the mark or whether

Machiavelli's advice was more appropriate for Renaissance princes than contemporary undergraduates.

The overall batting average indicated that there was a certain amount of internal consistency in Machiavelli's writings. It was surprising that speculations on the nature of man's behavior and how to cope with it made by a northern Italian more than four centuries ago still had enough cogency to divide even such haphazard samples as ours on items suggested by his writings. His astuteness is emphasized by the fact that, in general, the items adopted directly from Machiavelli were more discriminating than those we had contrived. For example, item 65, "It is a good working policy to keep on good terms with everyone," had been made up, but failed to discriminate because the majority of the respondents tended to agree with it.

Prior to administering the scale to the groups of college students, the 71 items had been classified as falling into one of three substantive areas. Thirty-two had been classified *a priori* as being concerned with the nature of an individual's interpersonal tactics, e.g., "The best way to handle people is to tell them what they want to hear" or a reversal, "One should take action only when sure it is morally right." In a second classification were the 28 items which appeared to deal with views of human nature, e.g., "Most men forget more easily the death of their father than the loss of their property" or a reversal, "Most people are basically good and kind." Eleven statements dealt with what might be called abstract or generalized morality, e.g., "People suffering from incurable diseases should have the choice of being put painlessly to death" and a reversal, "The world would be in much better shape than it is if people acted upon basic ethical principles." The fewest items appear in this last category because the construction of items tended to follow Machiavelli's writings rather closely and Machiavelli was less concerned with abstractions and ethical judgments than with pragmatic advice.

The classification system occasionally led to the somewhat arbitrary placement of an item. For instance, the rejection of the item, "All in all, it is better to be humble and honest than important and dishonest," might indicate disdain for interpersonal tactics of humility or disbelief that the meek shall inherit the earth. The great majority of the items, however, could be fairly easily assigned to one of the three categories.

Part-whole correlations were run between individual items and the subscales to which they had been arbitrarily assigned. Since no major differences emerged from the comparison of the part-whole subscale correlations with the item and total scale correlations, it did not seem imperative at that time to do a factor analysis to determine whether these dimensions were in fact factorially independent. (This point will be reexamined in Appendix A.)

The next concern in the item analysis was the relative efficacy of pro- and anti-Machiavellian items. Table II-2 contains a summary of the initial classification of the items and the direction of wording. From the table it is clear that, in general, statements worded in the Machiavellian direction tend to be more discriminating than the reversals.

TABLE II-2
Proportion of Items Significant at the .05 Level
by Content and Direction of Wording

| Direction | Content | | | | | | Total | |
|-----------|---------|------|-------|------|----------|------|-------|------|
| | Tactics | | Views | | Morality | | | |
| | P | (N) | P | (N) | P | (N) | P | (N) |
| Positive | .81 | (21) | .80 | (15) | .50 | (6) | .74 | (42) |
| Reversed | .55 | (11) | .69 | (13) | .60 | (5) | .62 | (29) |
| Total | .72 | (32) | .75 | (28) | .55 | (11) | .70 | (71) |

Another aspect of the analysis which was thought provoking was the relative discriminatory power of the items classified as falling into one or another of the three substantive areas. Again, the table indicates that there does not appear to be any marked difference in discriminatory power between items assigned at face value to any one of the subscales. However, it requires relatively small correlations with *Ns* of this size to produce significance at the .05 level. If an arbitrary criterion of a part-whole correlation of .20 is set as a minimum, a definite ordering among the three subscales emerges. Seventy percent of the items dealing with views of human nature pass this hurdle, as do 53% of those dealing with interpersonal tactics; only 36% of those dealing with abstract morality survive.

A LIKERT-TYPE MACH SCALE (IV)

Given a large pool of items which discriminated between high and low scorers on the total scale, the next problem was to decide which items to use for further research. The final scale was intended for making group comparisons and for selecting subjects for research rather than for individual diagnosis. Since it was to be administered to large groups of respondents, frequently in conjunction with

other materials, a relatively short version was desirable. The decision was to use 20 items on the assumption that these would give gross but sufficient discrimination in future samples without requiring an undue amount of time filling out scales by each respondent. Ten items were selected in which agreement was keyed to endorsement of Machiavellian statements and ten keyed in the opposite direction. This counterbalancing was designed to minimize the effects of indiscriminate agreement or disagreement with items.

Two additional considerations also guided the choice of items. First, it was advisable to select discriminating items reflecting a variety of content. A perusal of Table II-1 indicates that in some cases there were statements with similar meaning but phrased differently, and an attempt was made to avoid this in a standardized scale. At the same time items of greater discriminatory power were selected.

The items in Table II-3 were those finally selected after taking into account counterbalancing of wording, variety of content, and discriminatory power. This 20-item Likert format scale was named "Mach IV." (A version used before the item analysis was completed had been identified as "Mach III.") The mean item-whole correlation of these items was .38. Breaking these down by content area, the mean item-whole correlation for the nine items classified as dealing with Tactics was .41, for the nine on Views of Human Nature, .35, and for the two on Abstract Morality, .38. The mean part-whole correlation of those items worded in agreement with Machiavelli was .38; that of the reversals was .37.

The magnitude of this mean part-whole correlation was a delightful surprise. A bit of background might indicate why. The *F* scale developed by the authors of *The Authoritarian Personality* went through the laborious procedure of successive construction of items, item analysis, revision of items, and the writing of new ones. The third revision (in Forms 40 and 45) was subjected to an item-whole correlation on a sample of 517 undergraduate women. The mean correlation was .33 (Adorno *et al.*, 1950, p. 261). Since female respondents are typically less consistent than males on such scales, this might partially account for a lower correlation than those found among our respondents. On the other hand, the *F* scale correlations were probably enhanced because the items were all worded so that agreement indicated authoritarianism. These scores actually represented an additive function of both item content and response set.

Be this as it may, the first nine samples tested on Mach IV had a mean split-half reliability of .79; the same respondents had a reliability of .68 on the revision of the *F* scale by Christie, Havel, and Seidenberg (1958), which also had 20 items and had been counterbalanced for response set.

There is a possible explanation for this discrepancy. Most of the *F*-scale items were based upon an interpretation of Freudian theory; Mach-scale items were based upon the writings of Machiavelli. Freud's writings stressed the irrational components in human behavior and therefore dealt with complex, abstract

TABLE II-3

Item-Whole Phi Coefficients for 20 Most Discriminating
Mach Items Ordered in Terms of Mean Discriminability

| Item No. Mach | II | IV | Ident. | Item | Phi | | | |
|------------------|----|------------|--------|--|---------|------|------|----|
| | | | | | Hofstra | Iowa | N.C. | X |
| 45 | 2 | Tactics + | | The best way to handle people is to tell them what they want to hear. | 48 | 45 | 48 | 47 |
| 59 | 10 | Tactics - | | When you ask someone to do something for you, it is best to give the real reasons for wanting it rather than giving reasons which might carry more weight. | 55 | 50 | 31 | 45 |
| 11 | 12 | Tactics + | | Anyone who completely trusts anyone else is asking for trouble. | 41 | 38 | 48 | 42 |
| 30 | 18 | Views + | | It is hard to get ahead without cutting corners here and there. | 38 | 43 | 46 | 42 |
| 47 | 6 | Tactics - | | Honesty is the best policy in all cases. | 54 | 48 | 25 | 42 |
| 21 | 5 | Views + | | It is safest to assume that all people have a vicious streak and it will come out when they are given a chance. | 33 | 43 | 47 | 41 |
| 67 | 1 | Tactics + | | Never tell anyone the real reason you did something unless it is useful to do so. | 35 | 49 | 39 | 41 |
| 25 | 3 | Tactics - | | One should take action only when sure it is morally right. | 38 | 43 | 43 | 41 |
| 29 | 15 | Tactics + | | It is wise to flatter important people. | 41 | 37 | 43 | 40 |
| 70 | 9 | Morality - | | All in all, it is better to be humble and honest than important and dishonest. | 48 | 29 | 38 | 38 |
| 51 | 17 | Views - | | Barnum was very wrong when he said there's a sucker born every minute. | 43 | 30 | 39 | 37 |
| 31 | 19 | Morality + | | People suffering from incurable diseases should have the choice of being put painlessly to death. | 46 | 35 | 30 | 37 |
| 60 | 16 | Tactics - | | It is possible to be good in all respects. | 43 | 34 | 32 | 37 |

TABLE II-3 (continued)

| Item No. | | | Mach | Item | Phi | | | | |
|----------|----|-----------|------|---|-------------|------|------|------|------|
| II | IV | Ident. | | | Hofstra | Iowa | N.C. | X | |
| 53 | 4 | Views - | | Most people are basically good and kind. | 28 | 39 | 38 | 35 | |
| 68 | 7 | Tactics - | | There is no excuse for lying to someone else. | 35 | 37 | 32 | 35 | |
| 63 | 20 | Views + | | Most men forget more easily the death of their father than the loss of their property. | 34 | 28 | 39 | 35 | |
| 19 | 11 | Views - | | Most people who get ahead in the world lead clean, moral lives. | 34 | 34 | 30 | 33 | |
| 28 | 8 | Views + | | Generally speaking, men won't work hard unless they're forced to do so. | 30 | 32 | 30 | 31 | |
| 46 | 13 | Views + | | The biggest difference between most criminals and other people is that criminals are stupid enough to get caught. | 27 | 34 | 31 | 31 | |
| 44 | 14 | Views - | | Most men are brave. | 32 | 35 | 31 | 29 | |
| | | | | | $\bar{M} =$ | 39.2 | 37.7 | 37.0 | 37.9 |

concepts; Machiavelli was the realist par excellence. Since most scale construction is based upon the premise that people are at least rational enough to be consistent in answering questions, items based upon Machiavelli have a built-in advantage over those based upon Freud. This, of course, has nothing to do with the validity of the two theories. Machiavelli in dealing with overt behavior was more concerned with the more easily specified "how" than the more ambiguous "why."

A FORCED-CHOICE VERSION OF THE MACH SCALE (V)

Although it was initially gratifying to have a scale counterbalanced for response set which did not correlate with any known measure of agreement response set, one problem had not been anticipated. At the time Mach IV was being constructed, Edwards (1957) was pointing out the effects of social desirability upon responses to personality and attitude inventories. One sample

of medical school students who had filled out Mach IV had also been given the Edward's scale of Social Desirability. The correlation between the Edward's SD scale and Mach IV were $-.35$ and $-.45$ for the two classes included in both studies.¹ Budner (1962) found a correlation in the neighborhood of $-.35$ between the Edwards SD scale and Mach IV among college males but among college females this was inflated to roughly $-.75$. Since the latter correlation was almost as high as the reliability of Mach IV, it suggests that these female respondents were answering Mach items almost exclusively in terms of social desirability as defined by Edwards. Evidently the social norms of female undergraduates do not sanction the admission of interpersonal deceit when answering questionnaires.

In an attempt to bypass the effects of social desirability upon responses to the Mach IV scale, a new strategy was devised. In an earlier effort to obtain a less transparent measure of anxiety, the forced-choice format used by Heineman (1953) had been adapted for the Taylor Scale of Manifest Anxiety (Christie & Budnitsky, 1957). [Heineman had adapted the format from Naomi Stewart (1945) who had used it in personnel research in World War II.]

Many forced-choice formats require the respondent to indicate which of two equally repulsive alternatives is most true of him. Quite understandably, many respondents object. (One medical student taking such a test wrote on the margin, "This is like asking me whether I would rather rape my mother or take an axe to my father.") Stewart's technique was more subtle. Three instead of two items were presented with the instruction to choose which one of the three was most characteristic of the respondent and which one was least characteristic; thus, no response was marked for the third item. One of the three items was keyed for the particular scale of interest. It was matched in rated social desirability with another item known to be unrelated to the dimension under scrutiny. The third or buffer item was high in social desirability if the keyed and matched items were low. If the keyed and matched items were high in social desirability, then the buffer was low. Phenomenologically, in the first case the respondent was faced with an innocuous buffer item which most respondents (over two-thirds in most samples with which we are familiar) found little apparent difficulty in saying was most like them. The choice was then which of the two noxious items was *least like* them. If the buffer was low in social desirability, the respondent was then faced with the choice of saying which of two socially desirable statements was *most like* him.

This technique has a further virtue. It makes it difficult for the average respondent to determine which is the socially "correct" answer between the

¹Dr. Helen H. Gee, then the director of Research for the American Association of Medical Colleges, kindly made the Edwards' Social Desirability score available for the respondents in our mutual samples.

keyed and matched items. There is apparently an infinite regress in the competition between test constructors and takers. As the latter become more exposed to tests and more sophisticated in taking them, they find it easier and easier to divine what the tester is driving at and to give the responses they believe appropriate in the situation. [In an appendix to *The Organization Man*, Whyte (1956) attempted to out-“psych” the testers by giving his reader advice on how to cheat on personality tests.]

In constructing the Mach V scale we started with the 20 items in Mach IV and another, larger pool of items from scales known to be unrelated to it and which had a wide range of social desirability ratings. All items were randomly ordered and presented to respondents. Instructions to the raters asked them to indicate for each item how socially desirable someone whose opinion they valued (such as friends, teachers, or parents) would rate their endorsement of that item. If such endorsement would be viewed as highly undesirable, the item was to be given a rating of 1. If it would be viewed as highly desirable, it was to be given a rating of 5. Intermediate degrees of social desirability were assigned the intermediate ratings, with three being the point of indifference.

Two samples of raters were used. Previous experience with Heineman's forced-choice version of the Taylor Scale of Manifest Anxiety led to this decision. Heineman had used a sample of undergraduates at the State University of Iowa as judges. When we gave his scale to some fairly sophisticated seniors at the Western Reserve School of Medicine, there were certain incongruous findings. One of Heineman's triads was:

- A. I cannot keep my mind on one thing.
- B. I feel sure that there is only one true religion.
- C. Sexual things disgust me.

In this triad C is the negative buffer. A is the keyed anxiety item which was matched in social desirability (high) with B in the original Iowa sample of judges. About 90% of the Western Reserve medical students indicated that A was more like them than B. Since medical students were not more anxious than undergraduates, as measured by summed scores on the whole scale, the expectation would be that the two matched items should break about 50-50. One interpretation of this disproportionate choice of A, therefore, was that the items were not matched in social desirability for the more sophisticated medical students. They evidently did not ascribe as high a social desirability value to the “matched” statement, “I feel sure that there is only one true religion,” as did Iowa undergraduates.

At the time we were working with samples of medical school students and the Mach V scale was intended primarily for them. Their Mach IV scores were higher than those of college undergraduates at that time but not as high as those

of graduate students in psychology. We therefore had ratings of the social desirability of items made on a five-point scale by a sample of undergraduates in a not particularly elite college and by graduate students in a course in social psychology. Our suspicions were correct. The graduate students did give slightly but statistically significantly higher social desirability ratings to the Mach items than did the undergraduates. We therefore selected an item to match a particular Mach item if it had a rating within .2 of the Mach item ratings in *both* samples even if there was a marked difference in *mean* social desirability rating of the matched pair between the two samples. If samples both lower and higher in Mach scores than the medical students agreed on the desirability matchings, we assumed that the medical students were bracketed and that the items would be matched for them as well.

Although there were differences in the mean social desirability ratings by the two groups, the resulting ranking of the items was relatively consistent. The rank order correlation between the ratings of the 20 Mach items by undergraduates and graduate students was .89.

An attempt was made to have the buffer as far removed in social desirability value from the matched items as possible. Table II-4 indicates the triads finally assembled and the rated social desirability of the items within them (the mean is for judgments of both rating groups.)

The mean social desirability score is slightly unfavorable for the items keyed in the Machiavellian direction (2.63) and slightly on the favorable side (3.26) for the reversals. The difference in rankings is significant at the .05 level by the Mann-Whitney *U* test. The overall social desirability mean of 2.94 does not differ significantly from the theoretical neutral point of 3.00.

In the original scoring of the Mach V scale, the comparison was between the Mach items and the matched items. The respondent was required to say which item was most like him and which was least like him. On item 1 in Table II-4, for example, a point was scored if C (Mach item) was most like and A (matched) was least like, if C was most like and A was omitted, or if B (buffer) was most like and A was least like (this gave C which was omitted a higher rank than B). This procedure was followed when the Mach item was worded in the Machiavellian direction. When a Mach reversal was in the triad, as in item 5, a point was scored if B (matched) was most like and A (Mach) was either least like or omitted, or if B was omitted and A was least like. In practice it has been found that this intricate scoring system and the hidden nature of the forced choice makes it difficult for the average respondent to decide what the "right" answer is.

This conclusion is partially based upon the fact that relative to the neutral point on Mach IV and Mach V, some thousands of respondents consistently scored higher on Mach V (see Table II-6 as an example). Shortly after Mach V was developed, an advanced graduate sociology class in methodology taught by

TABLE II-4
Mach V
Mean Ratings of Social Desirability

| | Items | Mean ratings ^a |
|----|---|---------------------------|
| 1. | A. It takes more imagination to be a successful criminal than a successful business man. B. The phrase, "the road to hell is paved with good intentions" contains a lot of truth. C. Most men forget more easily the death of their father than the loss of their property. | 1.95 3.60 1.95 M |
| 2. | A. Men are more concerned with the car they drive than with the clothes their wives wear. B. It is very important that imagination and creativity in children be cultivated. C. People suffering from incurable diseases should have the choice of being put painlessly to death. | 2.85 3.70 2.85 M |
| 3. | A. Never tell anyone the real reason you did something unless it is useful to do so. B. The well-being of the individual is the goal that should be worked for before anything else. C. Since most people don't know what they want, it is only reasonable for ambitious people to talk them into doing things. | 2.4 M 3.7 2.35 |
| 4. | A. People are getting so lazy and self-indulgent that is is bad for our country. B. The best way to handle people is to tell them what they want to hear. C. It would be a good thing if people were kinder to others less fortunate than themselves. | 2.80 2.80 M 4.35 |
| 5. | A. Most people are basically good and kind. B. The best criteria for a wife or husband is compatibility — other characteristics are nice but not essential. C. Only after a man has gotten what he wants from life should he concern himself with the injustices in the world. | 3.60 M- 3.50 2.15 |
| 6. | A. Most people who get ahead in the world lead clean, moral lives. B. Any man worth his salt shouldn't be blamed for putting his career above his family. C. People would be better off if they were concerned less with how to do things and more with what to do. | 2.90 M- 1.80 2.95 |

TABLE II-4 (continued)

| | Items | Mean ratings ^a |
|-----|---|---------------------------|
| 7. | A. A good teacher is one who points out unanswered questions rather than gives explicit answers. B. When you ask someone to do something, it is best to give the real reasons for wanting it rather than giving reasons which might carry more weight. C. A person's job is the best single guide as to the sort of person he is. | 3.85 3.90 M- 2.30 |
| 8. | A. The construction of such monumental works as the Egyptian pyramids was worth the enslavement of the workers who built them. B. Once a way of handling problems has been worked out it is best to stick to it. C. One should take action only when sure it is morally right. | 1.70 3.45 3.45 M- |
| 9. | A. The world would be a much better place to live in if people would let the future take care of itself and concern themselves only with enjoying the present. B. It is wise to flatter important people. C. Once a decision has been made, it is best to keep changing it as new circumstances arise. | 2.1 2.9 M 2.75 |
| 10. | A. It is a good policy to act as if you are doing the things you do because you have no other choice. B. The biggest difference between most criminals and other people is that criminals are stupid enough to get caught. C. Even the most hardened and vicious criminal has a spark of decency somewhere within him. | 2.15 2.05 M 3.8 |
| 11. | A. All in all, it is better to be humble and honest than to be important and dishonest. B. A man who is able and willing to work hard has a good chance of succeeding in whatever he wants to do. C. If a thing does not help us in our daily lives, it isn't very important. | 3.85 M- 4.00 2.25 |
| 12. | A. A person shouldn't be punished for breaking a law that he thinks is unreasonable. B. Too many criminals are not punished for their crimes. C. There is no excuse for lying to someone else. | 2.15 3.10 3.15 M- |

TABLE II-4 (continued)

| | Items | Mean ratings ^a |
|-----|--|---------------------------------------|
| 13. | A. Generally speaking, men won't work hard unless they are forced to do so. B. Every person is entitled to a second chance, even after he commits a serious mistake. C. People who can't make up their minds are not worth bothering about. | 2.50 M |
| 14. | A. A man's first responsibility is to his wife, not his mother. B. Most men are brave. C. It's best to pick friends that are intellectually stim- ulating rather than ones it is comfortable to be around. | 4.0 2.75 M- 2.85 |
| 15. | A. There are very few people in the world worth con- cerning oneself about. B. It is hard to get ahead without cutting corners here and there. C. A capable person motivated for his own gain is more useful to society than a well-meaning but ineffective one. | 1.90 3.05 M |
| 16. | A. It is best to give others the impression that you can change your mind easily. B. It is a good working policy to keep on good terms with everyone. C. Honesty is the best policy in all cases. | 2.35 3.70 3.65 M- |
| 17. | A. It is possible to be good in all respects. B. To help oneself is good; to help others even better. C. War and threats of war are unchangeable facts of human life. | 2.55 M- 4.00 2.50 |
| 18. | A. Barnum was probably right when he said that there's at least one sucker born every minute. B. Life is pretty dull unless one deliberately stirs up some excitement. C. Most people would be better off if they control their emotions. | 3.55 M ^{a,b} 2.10 3.40 |

TABLE II-4 (continued)

| | Items | Mean ratings ^a |
|-----|--|---------------------------|
| 19. | A. Sensitivity to the feelings of others is worth more than poise in social situations. B. The ideal society is one where everybody knows his place and accepts it. C. It is safest to assume that all people have a vicious streak and it will come out when they are given a chance. | 3.8 2.25 2.20 M |
| 20. | A. People who talk about abstract problems usually don't know what they are talking about. B. Anyone who completely trusts anyone else is asking for trouble. C. It is essential for the functioning of a democracy that everyone vote. | 2.80 2.75 M 4.1 |

^aM indicates Mach original; M- indicates a reversal.

^bThis item was revised from No. 51 in Table II-1 and II-3 because some high-Mach respondents pointed out the original wording underestimates the number of suckers born, but the Mach IV item still reads "wrong" and this makes 11 + and 9 - Mach-worded items in Mach IV.

Paul F. Lazarsfeld was told the principle underlying the scoring method but was only told that the test was designed to measure agreement with Machiavelli. These students could not identify the keyed items.

More recently Marks and Lindsay (1966) have raised questions about the fakability of scores and noted that they could obtain highly significant shifts in Mach V scores from a standard administration to one in which "the subjects were provided with an extensive description of the interpersonal attitudes and behaviors of the so-called Machiavellian personality. Each subject was then instructed to respond to each triad of items as 'if he were a Machiavellian' [p. 230]." This is somewhat similar to having subjects take the TAT, having it scored for *n* Achievement, then telling the subjects the scoring system in terms of how striving and success themes in the stories are scored, and then telling them to take the test as if they were high on need for achievement. The crucial question, of course, as Singer pointed out [see Marks and Lindsay (1966, Footnote 9, p. 234)] is whether or not subjects taking the test under normal instructions can figure out what it is measuring and modify their responses accordingly. Singer (1968, personal communication) gave the scale in a variety of ways after a standard administration, for example: "Take it as if you wanted to make a good

impression on an employer"; "After reading the Appendix in Whyte's *The Organization Man* (1956) on how to fake on a test, fake low on this test"; and more interestingly, "Read the scale, decide what it means, and then make a high score on it." None of these procedures yielded Mach scores which differed significantly from those obtained using the standard instructions.

THE INTERPRETATION OF MACH SCORES

In all of the studies reported in this book the Mach scales have been given as a normal part of a research procedure. Usually, although not always, they have been given in conjunction with other items in a different context, by a different person than the experimenter, and up to six months prior to the actual experimental situation. If high Machs "saw through" the test, they were extremely cooperative in giving responses to the scales which were in high congruence with their predicted behavior in subsequent experimental situations. This interpretation admittedly violates some persons' notions of Machiavellians if the latter are viewed as so tricky that they will not answer a questionnaire honestly. Sceptics are invited to scrutinize the following studies in which the scales have been used for classification of subjects as high or low in manipulative tendencies and explain why positive relationships between scale scores and manipulative behavior in the laboratory have been found. It is, of course, possible that true manipulators are so clever that they fake low on the scales and then deliberately get conned by other subjects in the laboratory. This appears highly unlikely but it does raise the interesting question that if these super Machs are so busy dissembling, *when* do they manipulate.

The question has also been raised about how the scale works when it is so complicated to score. Heineman's findings are relevant here. He not only had subjects rate anxiety items for social desirability but also had them take Taylor's scale in its original agree-disagree format. He discovered that those respondents who scored high in measured anxiety rated the individual anxiety items as more socially desirable than those who scored low on the total scale. It is important to note the social desirability ratings are *means* of group judgments and that some individuals rank items higher in social desirability than others. The implication of Heineman's findings, then, suggests that the two "matched" items are *not* exactly matched for social desirability except for respondents near the mean on the total scale. If a respondent has a high scale score, he will tend to rank the keyed item as higher in social desirability than the item equated with it by the judging group. This suggests that there are two reasons why a high scorer might pick the keyed item over the matched one. First, it is more likely to represent his true belief; second, if he chooses on the basis of social desirability, his built-in bias will make him opt for the keyed item.

In most samples the reliability of Mach V hovers in the .60's. At first glance this is not overly impressive, although it is high enough to separate sheep from goats in some experimental situations as will be seen in subsequent chapters. There is one point that should be noted, however. The elimination of both response set and social desirability tends to decrease scale reliabilities. If our concern had been to construct a scale with higher internal consistency, this could have been done easily. We were more interested in devising a scale which would make meaningful discriminations among individuals' behavior. For this reason an attempt was made to minimize the effects of such possibly extraneous variables as response set and social desirability.

At the time we were constructing the scales we had to choose between alternative strategies. One was to focus upon purifying them to maximize internal consistency. The other was to determine whether or not the imperfect scales we had would be adequate for research. The decision was not to worry about psychometric perfection, but to find out if the scales had any relevance to the respondent's behavior.

The 71 items in the initial pool contained in Mach II were scored on a 1-5 point scale from strongly agree to strongly disagree, with a neutral midpoint of 3. The 20 items selected for Mach IV were scored on a 7-point scale; strongly agree, somewhat agree, slightly agree, no opinion, slightly disagree, somewhat disagree, strongly disagree. If the items were worded in the Machiavellian direction, the scale values varied from 1 for strong disagreement to 7 for strong agreement. If they were worded in the opposite direction, the scoring was reversed. Omissions were prorated if there were no more than three. Protocols with four or more omissions were discarded.

To simplify interpretation a constant of 20 has been added to all scores reported in this book. If someone's responses are counterbalanced so that his total score is at the theoretical neutral point, his score is 100 (20 items \times item mean of 4.0 + 20). The maximum possible score is 160 based on strong agreement with the 10 items worded in the Machiavellian direction and strong disagreement with the 10 reversed items ($20 \times$ item mean of 7.0 + 20) and the minimum score is 40. If fewer than 20 items are answered, the score is prorated so that the minimum score is again 40, the theoretical neutral point is 100, and the maximum is 160. A guide to the meaning of the scores is given in Table II-4A.

Upon occasion we have used a *Tendency to Agree Score*. Here the total score for the 10 items worded in the negative direction is subtracted from that of the 10 items keyed positively and a constant of 100 is added. If a respondent slightly agrees with every item, regardless of content, he would obtain a score of 120 ($10 \times 5.0 - 10 \times 3.0 + 100 = 120$). If his agreements and disagreements balanced out, he would get a score of 100. Scores higher than 100 indicate tendencies to indiscriminate agreement; scores lower than 100 indicate a tendency to disagree.

TABLE II-4A

Mean Response Tendency

| Mean score | Pro items | Anti items |
|------------|-------------------|-------------------|
| 40 | Strongly disagree | Strongly agree |
| 60 | Somewhat disagree | Somewhat agree |
| 80 | Slightly disagree | Slightly agree |
| 100 | Neutral | Neutral |
| 120 | Slightly agree | Slightly disagree |
| 140 | Somewhat agree | Somewhat disagree |
| 160 | Strongly agree | Strongly disagree |

We have used this score primarily in a secondary fashion. Samples with means close to 100 on both Mach IV and Tendency to Agree obtain their scores by responding more to content than by indiscriminate agreement or disagreement. Samples which have Mach IV scores near 100 but whose Tendency to Agree scores vary markedly from 100 are answering more in terms of response set than the content of the items.

The original scoring on Mach V led to a range of zero to 20, 10 being the theoretical neutral point. This system kept the comparison between the Mach and matched items separate from the buffers. A subject, for example, could say every buffer was most or least like him. This still left 20 comparisons between the Mach items and the matched items so that the range of scores went from zero to 20. Interestingly enough, the correlation between Mach scores and the built-in scale of social desirability remained significantly negative.

A SUBSIDIARY MEASURE OF SOCIAL DESIRABILITY

A possibility was to use the buffer items on Mach V as yet another way of measuring social desirability. If, on those ten triads in which the buffer was higher in social desirability ratings than the two matched items, one point was allocated for every occasion in which a respondent said the buffer was most like him; no points were assigned in which it was checked as least like him or omitted in the comparison. Similarly, in the ten triads in which it was lower in social desirability than the matched pairs, a point was assigned when it was checked as least like the respondent and no points were assigned if it was checked as most like or omitted. Scores on this derived measure of social desirability thus could vary from zero to 20 just as Mach V scores.

Over two-thirds of the buffer items were appropriately checked as most or least like the respondent in every known sample, which indicates that the separation of the buffer from the two matched items in social desirability ratings was highly successful (chance expectation would be only one-third). Because of the relatively low variability in such a restricted range we have not found this derived measure of social desirability particularly useful as a measure of individual differences although we did find, as might be expected, that samples of respondents high in judged middle-class conventionality scored significantly higher.

There is one point which should be clarified since the correlation between this measure of Social Desirability and Mach scores is significantly negative. For example, in a heterogeneous aggregation of students in 14 different colleges (see Table II-6) the correlations were $-.47$ for 764 Caucasian males and $-.35$ for 832 Caucasian females. Negative correlations are built into the scale format. With the one-and-one scoring system for Mach V and Social Desirability the following matrix indicates why:

| | | Most Like | | |
|--|------------|------------|------|-----------|
| | | 1, 0 | 1, 0 | — |
| Response to keyed positive Mach item in Mach V | Omit | 0, 0 | — | 1, 1 |
| | Least Like | — | 0, 0 | 0, 1 |
| | | Least Like | Omit | Most Like |
| Response to high SD buffer item on Mach V | | | | |

The first number in each of the six possible pairs refers to the point counting toward a score on Mach V, the second to the score on Social Desirability. It is impossible to check both the keyed Mach item and the buffer item as most like, omit both, or as least like, respectively, so these combinations are crossed out. Looking at the pairings above the diagonal, it is clear that a predominance of Mach "Most Like" item endorsements over the buffer leads to a high score on Mach and a zero score on Social Desirability. If the buffer item is checked as "Most Like" by the respondent, it is possible for one of the two responses to be scored for Mach but not the other. A negative correlation between Mach and Social Desirability Scores is thus built into the scoring system.

If respondents answered randomly, we calculate that there would be a correlation in the vicinity of $-.55$ between Mach V and Social Desirability total scores. However, we know that respondents do not respond randomly and actually choose the buffer as most like them in over two-thirds of the cases. This reduces the expected built-in negative correlation to the vicinity of $-.40$. This means that we could expect the correlation between the two scales to be highly negative among samples high in Machiavellianism and lower but still significantly negative in samples which are relatively high in Social Desirability.

Singer (1964, p. 136) found a negative correlation of $-.35$ between Mach V and the internal measure of social desirability; Marks and Lindsay (1966) found one of $-.44$. In instances where *external* measures of Social Desirability, constructed by Edwards (1957) and Crowne and Marlowe (1960), have been used, no significant correlation has been found with Mach V.

Upon the basis of available evidence it seems reasonable to say that scores on Mach IV reflect Machiavellian orientations not only because respondents agree with Machiavelli but are also willing to endorse socially undesirable statements. Scores on Mach V reflect the willingness of respondents to agree with Machiavelli when their tendency to agree with social undesirable statements is removed.

An alternative scoring system for Mach V was developed with two considerations in mind: (a) to have the possible range and theoretical neutral point equivalent to Mach IV, and (b) to take full advantage of the fact that it is probably more Machiavellian to say the Mach item is most like and the matched item least like oneself — a two-step difference — than to say the Mach item is most like and omit the matched item or omit the Mach item and say the matched item is least like oneself — a one-step difference.

Under this system the following item scoring is used when the Mach item is worded in the pro direction:

| Mach item | Matched item | Score |
|------------|--------------|-------|
| Most Like | Least Like | 7 |
| Most Like | Omitted | 5 |
| Omitted | Least Like | |
| Omitted | Most Like | 3 |
| Least Like | Omitted | |
| Least Like | Most Like | 1 |

When the Mach item is worded in the anti direction, the scoring is also reversed, 1, 3, 5, 7, from top to bottom.