A COMPARISON OF TWO METHODS OF CONTENT

ANALYSIS ON STORY COMPLETIONS FROM THREE COUNTRIES

Thesis for the Degree of M. A.

MICHIGAN STATE UNIVERSITY

Kenneth W. Terhune

1961

ABSTRACT

A COMPARISON OF TWO METHODS
OF CONTENT ANALYSIS ON STORY COMPLETIONS
FROM THREE COUNTRIES


by Kenneth W. Terhune


This study compared the categorizing and rating scale
methods of content analysis applied to children's story completions
from Anderson's cross-national study. The purpose was to examine
empirically the merits and demerits of each technique for analyzing
qualitative data. As Anderson has employed the content coding
method in his analysis, rating scales were examined as a possible
alternate or complementary technique. Anticipated were the fol-
lowing advantages: (1) more molar assessment of story content;
(2) assessment of degree, rather than presence-absence; (3) applic-
ability across different Anderson stories; and (4) ability to ana-
lyze multifactorially, accounting for demographic variables ad-
ditional to nationality.

Comparison of the two methods was made on 301 completions
to Anderson's Damaged Axe story, in which a child unintentionally
damages his older brother's new axe. These completions were ran-
domly selected from the Anderson data, and were comprised about
equally of stories by American, Norwegian, and German children.

To analyze the stories by the content category method,
a coding manual of seventeen categories was devised. These
assessed mainly

(1)  The description of intent on the part of the younger
     brother in damaging the axe.

(2)  Handling of the situation by the older brother.

(3)  Reaction of the younger brother.

Inter-coder reliability of 92% was obtained with the coding method.
Differences across samples were tested with chi-square.

Kenneth W. Terhune

To analyze by the rating scale method, numerous scales were tried. Scales which were found unreliable or inapplicable to all stories were eliminated. In the final analysis, two scales were used:

(1) <u>Conflict Scale</u> - Rated the older brother's behavior on a conflict-harmony dimension.

(2) <u>Emphasis Scale</u> - Rated relative story emphasis on materialistic vs. interpersonal matters.

A simple analysis of variance was first made between the samples, and then a special multifactorial analysis for non-orthogonal data was employed. This method, which accounted for factors of nationality, sex, religion, and socio-economic status, is described in detail.

The content category results showed for the Norwegian stories a strong image of the older brother, for the American stories a strong emphasis on the parents, while German stories typically stressed punishment. The rating scales revealed most conflict in the German stories, and less respectively in the American and Norwegian samples. No nationality differences were found in relative story emphasis.

A major finding in the rating scale analysis was that significant sex, religion, and socio-economic differences exist in the data. Sex was equally as important as nationality on the Conflict Scale, while socio-economic status was the major factor on the Emphasis Scale. Because the demographic factors were found significant, the danger of confounding due to disproportionality was clearly demonstrated. Also discussed was interpretation of results in regard to causality and nationality differences.

It was concluded that rating scales did in general meet the original expectations. Their main advantage was considered to be ability to analyze multifactorially, which enabled (a) nationality differences to be determined <u>free</u> from effects of sex, religion, and socio-economic differences, and (b) the contribution

of nationality <u>relative</u> to the other factors to be shown. In the conclusion, the "content category" and "rating scale" methods were reconsidered as representing respectively attempts at nominal-ordering and interval-ordering of data categories. Implications of these approaches for subsequent research with qualitative data were discussed.

Approved _C F Wrigley_____
Signature of Major Professor
Date_July 15, 1961_____

A COMPARISON OF TWO METHODS

OF CONTENT ANALYSIS ON STORY COMPLETIONS

FROM THREE COUNTRIES


By

Kenneth W. Terhune


A THESIS


Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of


MASTER OF ARTS


Department of Psychology


1961

# ACKNOWLEDGEMENTS

TABLE OF CONTENTS

# LIST OF TABLES

LIST OF APPENDICES

# I.  INTRODUCTION

This study compares the categorizing and rating scale methods of content analysis applied to children's story completions from Anderson's cross-national study. The purpose is to examine empirically the merits and demerits of each technique for analyzing qualitative data. As Anderson has employed the content coding method in his analyses, rating scales will be considered as a possible alternate or complementary technique.

## The Anderson Cross-National Study

For assessing differences among countries, Anderson devised a projective technique employing eleven incomplete stories, in which situations of potential conflict involving children are described (see Appendix A). In completing these stories, a child is assumed to reveal his perceptions of social relations and some influences of his culture. Accordingly, these stories have been employed as a device for examining nationality differences among eight countries of Europe and North and South America.

The samples obtained were not necessarily representative of their respective countries nor of the particular cities from which they were selected. Rather, Anderson simply attempted to obtain groups of children from the several countries, each sample to contain boys and girls from the high, middle, and low socio-economic levels (Anderson, 1957).

In analyzing the children's story completions, content analysis has been achieved by empirically determining categories of information into which the responses may be placed. On this basis a coding manual is devised for each story, and coders are trained to check the appropriate categories for each child's story completions.

Frequencies of usage are then tabulated for the different categories on each location sample. From these data, significant differences in the use of the categories among the different samples are determined, using chi-square.

1

As of this writing, analyses have been made on stories 2
through 6 of Series A and stories 2 and 3 of Series B. Statistically
significant differences have repeatedly been found among the samples
on many of the response categories (Anderson & Anderson, 1956;
Anderson et al, 1957; Anderson et al, 1959; Anderson, 1960; Heber,
1955; Levitov, 1960; and Robinson, 1955).

## Additional Analyses Proposed

After reviewing the above studies, it seemed to the writer
that understanding of the Anderson data could be furthered through
the following additional analyses:

1. Examination of the general configuration of each
   child's stories. While the previous studies re-
   ported many differences on rather atomistic details
   within the story completions, it would be interesting
   to learn if the broader themes of the stories would
   also show such differences. Molar indices of content
   would facilitate such an analysis.

2. Assessment of each story as to degree rather than
   simple presence or absence of relevant variables.
   Such analysis would seem to be not only more precise,
   but more informative as well. A direct comparison of
   sample averages on unitary psychological dimensions
   could be made, and parametric statistics could be ap-
   plied when desired.

3. Comparison of sample responses to different incomplete
   stories. As each child writes completions to several
   different stories, a measuring device that could trans-
   cend specific content and be applied to different
   stories would help to generalize results.

4. Determination of contribution of variables other than
   location. A clearer indication of the nature of nation-
   ality differences can be obtained when effects of other
   variables are taken into account. This will prevent

misinterpretation from confounded variables and may re-
veal interaction between location and other variables
which should facilitate the understanding of national
character and nationality differences.

As the last point will be heavily emphasized, fuller con-
sideration is indicated. Indications from numerous sources suggest
the importance of certain demographic variables in cross-national re-
search. Included among these are some previous studies by the writer
on the Anderson data (Terhune, 1959a, 1959b, 1960). These studies
were restricted to responses to Anderson Incomplete Story A-3, and
results in only a few locations were examined. However, analysis
was limited by the inability of the chi-square simultaneously to
examine more than two variables on non-orthogonal[1] data. Neverthe-
less, results suggested that differences of sex, religion, and socio-
economic status, as well as location were influencing the stories.
Furthermore, these variables seemed to be interacting with each
other and with the variable of geographic location. Results thus
suggested the need for more detailed analysis of the influence of
these demographic factors.

Other studies indicating the importance of sex, religion,
and socio-economic status in cross-national research will next be
discussed.

a) _Sex_. It is common knowledge that cultures often as-
sign different roles to the two sexes (see, for example, the writings
of Margaret Mead). If a culture is characterized by an authori-
tarian family structure, these differences tend to emerge in the
form of masculine dominance and feminine submissiveness (Schaffner,
1948). Such differential sex roles have even been institutionalized
in Germany, for the civil code specifically states that the father
is head of the household, with ultimate authority on family decis-
ions (Lowie, 1954).

Empirical evidence of interaction between sex and location
is provided by McGranahan (1946), who found that girls in Germany

---

[1] Non-orthogonality exists when sub-class frequencies are dispro-
portionate.

and America tended to reflect their respective national value
patterns more than did the boys. Furthermore, German and American
girls tended to be at opposite extremes in their attitudes.

b) Socio-economic status. Rodnick (1955) states that
Germany has a very rigid class system, where everyone must "know
his place', and behave toward others accordingly. In contrast, the
class boundaries in the United States and Norway are less sharp,
as indicated by the subjects in the cross-national survey by
Buchanan and Cantril (1953). Similar observations were made by
German divinity students when visiting the United States (Lowie,
1954). In a study of sexual differentiation and identification,
Azimi (1960) found class and nationality differences between
Iranian and American children. Norwegian class differences were
noted by Rodnick (1955), while numerous social class studies in
the United States have shown class to be a factor associated with
personality differences (reviews of studies are given by Auld,
1952; Cattell, 1945; and Davidson, 1943).

c) Religion. Rokeach (1956) reports that significant
differences were found between Protestants and Catholics on re-
sponses to the California Fascism and Ethnocentrocism Scales, as
well as on his own Dogmatism and Opinionation Scales. Findings
were that Catholics were more "fascistic", "ethnocentric", "dog-
matic", and "opinionated"; it is hypothesized that these tendencies
would be also reflected in responses to the social conflict sit-
uations of the Anderson stories.

## Advantages of Rating Scales

Rating scales seem able to meet requirements specified
in the previous section. First, they should be readily applicable
to assessment of molar qualities, such as "harmony", "integration",
and "domination". Second, they by nature assess degree, rather
than presence-absence of a quality. Third, any generalized rating
scale should be applicable to different Anderson stories. And
fourth, by assuming approximately interval measurement, we may ap-
ply such statistical methods as analysis of variance or correlation
analysis to examine simultaneously the contributions of several

factors. In expectation of these advantages, we therefore propose to examine rating scale s as an alternate method to category analysis in assessing the Anderson stories.

## Objective of the Thesis

The research reported in this thesis will assess story completions by means of both category coding and rating scale methods. The relative merits of the two methods will then be compared, consideration given to the following methodological questions:

(1) What kinds of information can the two methods provide? Do rating scales indeed lend themselves more readily to assessment of molar psychological qualities than do content categories?

(2) Is there a difference in the cost required by the two methods in terms of time for development of the coding scheme, training of coders, and analysis of results?

(3) Are there differences in the inter-coder reliability of the two methods?

(4) What statistical analyses can be used with the two method s?

(5) Is one method superior to the other in showing differences between different samples?

## II. PROCEDURE

The writer chose to examine the story titled "The Damaged Axe", which is story 4 of Anderson Incomplete Stories, Series B. In its English version, Story B-4 is as presented below. For translation into other languages, the only alteration was that English names were replaced by boys' names common in the locations sampled.

### Story B-4: The Damaged Axe

Herbert received for his thirteenth birthday a handsome camping axe. It is sharp and has a strong leather case. While Herbert is at school his four-year-old brother Billy sees the axe, looks at it a long time, picks it up, puts it back, and finally takes it outdoors with him to play. Billy does not take the case off. He sings to himself as he walks about the garden, tapping the axe gently on a tree, a post, and the pavement. Herbert comes home from school, finds the axe in its leather case with some other toys. But he sees that the axe has cut through the leather case and the blade is chipped and blunted.

Following the story were the questions "What does Herbert do?" and "How does Herbert feel about it?".

### Psychology of the Story

Notice that the story does not describe any conflict, only a potential conflict situation. Motives, intentions, and outcome are not specified in the story, although there are implications that (1) Billy realizes that he ought not to touch the axe, and (2) he did not wittingly damage it. It is left to the child author to project the behavior and motivations into the story characters as he sees fit.

A child's responses may be expected to reveal his perception of social causality and his understanding of the behavior of a four-year-old child. Herbert's behavior may be described as more antagonistic and vindictive if Billy is perceived as an evildoer, whereas Herbert may act more constructively if Billy's behavior is seen as innocent, but lacking understanding of the

consequences of his actions.

The importance to the child of material possessions as opposed to interpersonal concerns may also be revealed. The damaged axe and its restoration may be the dominant theme of the story, or in contrast, emphasis may be on Billy's offense against Herbert and the reestablishment of harmonious relations.

Two examples may be given as to how the story may bring out differential effects of nationality, religion, sex, and socio-economic status. First, authoritarian tendencies exhibited by a national, religious, or socio-economic group may be reflected in more dominating, intolerant, hostile, and aggressive behavior on the part of Herbert. Second, it may be hypothesized that boys' stories will differ from girls' because they identify more with the characters in the story.

## Description of the Sample.

Anderson's sample. The parent population from which Anderson drew his samples is defined here as:

> A population which consists of children from the respective countries who are (a) attending school, (b) literate, (c) live in the specific cities sampled, (d) can write legibly and comprehendably, and (e) were born in the country from which the sample was taken.

In obtaining his subjects, Anderson did not attempt to get random samples. Instead, an effort was made to secure representation of the upper, middle, and low socio-economic levels. The choice of schools to insure coverage of these levels was made with the assistance of associates and school officials in the various locations. Primarily seventh-grade children were sampled, while additional grades were obtained in several locations for cross-grade analyses.

The story completions were obtained over a two-day period in all locations, generally on consecutive days. Series A was given on one day, and Series B on the other. Standardized instructions in the native tongue were given by assistants. (See Appendix C for instructions.)

Sampling for this study. For this research, a sub-sample
of the completions to story B-4 was selected. This was necessi-
tated by the considerable amount of work in copying, ordering, and
rating the stories. Three different countries were deemed suffic-
ient.

The sub-sample was taken from the responses of seventh-
grade children from Benton Harbor (Michigan), Drammen (Norway),
Munich (Germany), and Hamburg (Germany). American and German
samples were chosen because the two nations from which they were
taken have been commonly characterized as opposites on a demo-
cratic-authoritarian dimension; the German society has been imputed
to be the more authoritarian (Abrahamsen, 1945; Anderson et al,
1959; Fromm, 1941; Lowie, 1954; McGranahan, 1946; Schaffner, 1948).
Furthermore, these two groups have consistently exhibited consider-
able differences in previous analyses of the Anderson data (Anderson
& Anderson, 1956; Anderson et al, 1957; Anderson et al, 1959;
Anderson, 1960). Drammen was used because indications are that the
Norwegian society represents an intermediate position between the
United States and Germany, in relation to authoritarian tendencies.
The writer has heard this view expressed by Norwegian and Swedish
social scientists, and more objectively, previous analyses of the
Anderson data have shown the Norwegian to fall between the American
and German samples (Anderson, 1960).

About 100 children each were selected from the American,
Norwegian, and German samplings. The only continental United
States location in which story B-4 had been written was the Benton
Harbor sample, while Drammen was the only location available from
Norway. In Germany, B-4 stories had been written both by Hamburg
and by Munich children.[2] There were few Catholic children in Hamburg

---

[2].Story B-4 was also written in Braunschweig, Germany, but this
sample was obtained under special conditions, which did not al-
low comparison with the others.

while Protestants were equally scarce in Munich. Therefore, to secure adequate representation of both religions in the German sample, fifty children were chosen from each city.

Stories from the Anderson sample were selected by using the RAND Corporation list (1955) of random numbers. In a few cases where the story chosen was illegible, incomprehensible, or completely irrelevant, another random story was substituted.

The sex and religion of each child were obtained from the Anderson records, while socio-economic status of each child was determined on the basis of parents' occupations. (Father's income, place of employment, place of residence, and other standard indices of socio-economic status were not available.) The occupations were specified by the children, and were translated for this study.[3] Miller and Swanson's (1958) method was employed to index socio-economic status. However, the method was not strictly adhered to for every child, for the translators occasionally provided information which indicated a certain occupation to fall into a different socio-economic level in Germany or Norway than indicated for this country. (See Appendix D for more details.)

The breakdown of the sample selected for this study is shown in table 1. Notice that the proportions on the different factors vary greatly. Particularly notice the absence of any Drammen Catholics,[4] Catholic boys in Hamburg, and near-absence of Protestant boys in Munich.

## Analysis by Content Coding Categories

For coding the various types of information provided by the story completions, a coding manual was needed which would be

---

[3] Mr. John Hornslien and Dr. Victor Krampl translated and helped to determine the socio-economic levels. Mr. Hornslien, who translated the Drammen occupations, is a Ph.D. candidate in sociology and a native of Norway. The German occupations were translated by Dr. Krampl , a biochemist. He is a native of Czechoslavakia and fluent in German.

[4] While the Anderson records do not specifically state each child's religion in the Drammen sample, the population of Norway is reported to be almost homogeneously Lutheran (Rodnick, 1955).

Table 1

Composition of the Sample

(Numbers given are frequencies)

| | | U.S.A. (Benton Harbor) | | Norway (Drammen) | | Germany | | | | | |
| | | | | | | Hamburg | | Munich | | Total | |
| | | Boys | Girls | Boys | Girls | Boys | Girls | Boys | Girls | Boys | Girls |
|---|---|---|---|---|---|---|---|---|---|---|---|
| S.E.S.-1 (Upper middle) | Protestants | 3 | 2 | 10 | 8 | 4 | 5 | 0 | 1 | 4 | 6 |
| | Catholics | 4 | 2 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 6 |
| | Other | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| S.E.S.-2 (Lower middle) | Protestants | 6 | 10 | 21 | 16 | 7 | 12 | 0 | 3 | 7 | 15 |
| | Catholics | 4 | 5 | 0 | 0 | 0 | 1 | 6 | 14 | 6 | 15 |
| | Other | 2 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 2 |
| S.E.S.-3 (Lower) | Protestants | 15 | 21 | 28 | 19 | 11 | 5 | 1 | 3 | 12 | 8 |
| | Catholics | 6 | 3 | 0 | 0 | 0 | 2 | 4 | 11 | 4 | 13 |
| | Other | 8 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Prot., S.E.S. level unknown | | 3 | 3 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| Sub Totals | | 52 | 48 | 59 | 43 | 22 | 28 | 11 | 38 | 33 | 66 |
| Grand Totals | | 100 | | 102 | | 50 | | 49 | | 99 | |

Summary Data across Sub-Samples:

| | Boys | Girls | Prot. | Cath. | SES-1 | SES-2 | SES-3 |
|---|---|---|---|---|---|---|---|
| U.S.A. | 52 | 48 | 57 | 24 | 12 | 27 | 55 |
| Norway | 59 | 43 | 102 | 0 | 18 | 37 | 47 |
| Hamburg | 22 | 28 | 44 | 4 | 9 | 22 | 18 |
| Munich | 11 | 38 | 8 | 41 | 7 | 23 | 19 |
| Total German | 33 | 66 | 52 | 45 | 16 | 45 | 37 |

S.E.S. = Socio-economic Status

comprehensive and also permit high inter-coder reliability. Three
areas were considered essential for assessment:

(1) The description of intent on the part of the younger
brother in damaging the axe.

(2) Handling of the situation by the older brother.

(3) Reactions of the younger brother.

To determine the optimal organization of the specific
categories of information, several hundred stories from all lo-
cations were read. When interesting details appeared which the
writer had not previously anticipated, these were introduced into
the coding scheme, provided that such details seemed of psycholog-
ical importance and substantial frequencies were indicated. After
the first draft, the coding scheme was gradually improved in a
series of trial codings and reliability checks. Definitions were
modified, and categories with low frequency or low reliability were
eliminated. Whenever possible, categories were devised with mutually
exclusive and mutually exhaustive items. The coding recommendations
of Festinger and Katz (1953) and Selltiz et al (1959) were used
throughout for guidance.

Inter-coder reliability was determined by means of the
following standard formula (Festinger & Katz, 1953, p. 411):

$$\text{Reliability} = \frac{2x \ (\text{No. agreements between coder A and coder B})}{\text{No. items coded by A + No. items coded by B}}$$

The writer's codings were compared with those of another psychology
graduate student, who used only the coding manual for guidance.
Through the series of improvements mentioned above, reliability was
gradually increased from an initial 86% to a final figure of 92%.
At this point, the manual was considered satisfactory. Some 200
man-hours are estimated to have been spent in preparing the manual.
An outline of the final coding scheme is presented in Appendix E.

All stories in the research sample were coded by the
writer, at an average rate of about twenty stories per hour. The
codings for each story were punched on an I.B.M. card, and coding
frequencies for the three locations were calculated on the I.B.M.
number 101 Electronic Statistical Machine. Chi-squares were then

calculated for each category.

## Analysis by Rating Scales

First proposals. Initial formulation of dimensions to be
assessed by rating scales was accomplished by (a) determining the
psychologically important characteristics of the story that could
be assessed only by rating scales and (b) reviewing the content
categories for aspects that could also be scored on rating scales.
The writer was interested first in whether focus of the story cen-
tered on the damaged axe and its restoration, or more on the inter-
personal offense and the handling of the offender. This involved
a dimension not only of relative emphasis, but whether the behavior
described was essentially constructive or destructive. Therefore
the first scales attempted were three in number: (1) quality of
interpersonal relations (harmonious-discordant), (2) degree of
material accomplishment (constructive or destructive actions) and
(3) relative emphasis on interpersonal vs. materialistic aspects.
The first of these had to be rejected because its generality and
vagueness were found to render confident rating virtually impos-
sible. The second scale was eliminated because (a) behavior re-
garding the axe could more readily be placed in discrete descrip-
tive categories than on a continuum and (b) destructive behavior
toward the axe was seldom exhibited. Thus, only the emphasis
scale appeared practicable in this area. A second area of inter-
est related to authoritarian relationships, and, more specifically,
behavior of the story characters on a dominance-submission dimen-
sion. Accordingly, dominance-submission scales were established
for Herbert (the older brother) and Billy (the younger brother).
Trial ratings of stories on these scales revealed that the second
was worthless as an assessment tool, because the younger brother's
responses were seldom given in the stories, and even then his be-
havior usually was insufficiently described.

At this point the coding manual was again reviewed for
rating scale possibilities. The following are the rating scales
that resulted:

1. <u>Dominance Scale.</u> Herbert's dominance-submission in regard to Billy.

2. <u>Self-Sufficiency Scale.</u> Herbert's self-sufficiency in regard to restoring the axe.

3. <u>Conflict Scale.</u> Direction of Herbert's behavior: conflict-harmony.

4. <u>Emphasis Scale.</u> Emphasis in story: material vs. interpersonal.

5. <u>Punishment Scale.</u> Severity of punishment for Billy.

6. <u>Closure Scale.</u> Degree of closure to the story plot.

7. <u>Final Mood Scale.</u> Ending of story: optimistic-pessimistic.

Note that some of these scales pertain to specific aspects subsumed under the formerly proposed more global scales. It was found, for example, easier to rate Herbert's behavior in a dimension of conflict-harmony than on the more vague and general "quality of interpersonal relations".

<u>Pilot Study.</u> For a pilot study, each scale was presented on a form, with each continuum separated into nine divisions. As it was found quite difficult to provide differentiating phrases for cues along the scales, the writer chose to define only the midpoints and end positions on each scale. This also tended to prevent the writer's bias from affecting the judges' decisions as to where on a scale certain behavior should be placed. Brief instructions were written to provide some guidance to the judges.

Three graduate students of psychology, including the writer, were used for the pilot study rating. Twelve stories each from Hamburg and Munich, plus twenty-four stories from Benton Harbor were rated. (None of these stories were used in the main study.)

Of the seven scales, three were found unsatisfactory for the following reasons:

    <u>Self-Sufficiency Scale:</u> Only a minority of stories could be rated on this scale, because Herbert was frequently described as taking no action in regard to the axe.

Closure Scale: Low inter-coder reliability (about .54
between two raters), and no apparent differences be-
tween locations.

Final Mood Scale: Low inter-coder reliability; for all
samples, ratings averaged close to the mid-point of
the scale.

With certain improvements, the four remaining scales ap-
peared promising. The Dominance Scale was revised to exclude the
submission aspect,[5] while the Emphasis Scale reliability (about
.47) needed to be increased. To raise reliability on all the
scales, three procedures were applied. These were (a) adding cues
wherever possible, (b) providing sample stories at different points
along each scale, and (c) having a short training period for the
judges.

Cues for the rating scales were determined by selecting
stories in which the three judges had given nearly equal ratings
in the pilot study.[6] For the Conflict and Punishment Scales it

---

[5] At the suggestion of Dr. Harold H. Anderson, the meaning of the
original scale was re-examined and the conclusion drawn that "sub-
mission", meaning "non-resistance", was not a scalar variable.
One either submits to certain demands or one does not. It may
be possible to form a Guttmann scale of demands, to which one
stops submitting at a certain point, but, by our definition, sub-
mission itself cannot be so scaled. Dominance, on the other
hand, we do see as variable. Dominance varies as to the demands
one makes upon another, and as to the efforts one makes to con-
trol the behavior of another.

[6] Admittedly, there is danger of unreliability in choosing stories
agreed upon by only three judges. The chance factor is more
likely to be of influence, and bias may result because of the
particular judges used. In a repetition of this scale-develop-
ment procedure, more judges would be preferable, in order to de-
termine representative stories for different points along the
scales. Despite these admitted limitations, however, we still
have gained the advantage of providing cues which will function
as anchor points for use of the scales, thus increasing their
reliability.

was possible to determine what specific behavior elicited a certain rating, and therefore descriptions of these behaviors were used as cues along the scales. It was not possible to do this with the Emphasis and Dominance Scales. For all scales, however, stories on which judges agreed were selected from the pilot study as examples for the different scale positions. These, along with cues, were then included in the revised version of the written instructions to the judges (Appendix F).

Because some raters felt a need for more precision on the scales, the number of divisions was increased from nine to eleven. The final form of the rating scales is presented in Appendix G.

Main analysis. For the main study, three raters again were employed. All were psychology graduate students. One had participated in the pilot study, but the other two were not familiar with the Anderson research. To avoid bias, the writer ruled himself out as a rater.

The training for the raters consisted of a four-hour session in which thirty-six stories were rated and discussed. (None of these stories were used in the main study.) As a result of this discussion, supplementary instructions (Appendix F) were issued to clarify certain points.

Stories in the main sample were then prepared for rating. To avoid bias, all stories were copied with no identification of their source. The boys' names were given throughout as Herbert and Billy, and too-literal translations were converted to the appropriate English idiom. (Changes were conservatively made, and where the intended meaning was doubtful, the stories were left as translated.) If money was mentioned, appropriate American values were substituted. Because the spelling and punctuation of the Norwegian and German stories benefited from the translation process, the spelling and punctuation of the American stories were also corrected. Grammatical errors were left unaltered, because such errors were also retained in the translations.

Unknown to each judge, the first fifty stories he rated were warm-up stories. The ratings for these were deleted from the main

study, as it was assumed that the judges' adaptation levels were unstable during their initial ratings. To prevent any subsequent drift in adaptation level from affecting one location more than another, the rest of the stories were presented in random order. The judges were instructed to rate the stories only in the given order, and each was presented with a different order.

About eight to ten actual working hours were required by each judge to complete the ratings of the 50 warm-up stories and the 301 sub-sample stories. While no numbers were used by the raters on the rating forms, each scale division was later assigned a number from 1 to 11 or 1 to 5 (the latter for the Emphasis Scale). Each judges' ratings were thus recorded numerically in compiling the data, and the ratings were averaged for every story.

To check inter-judge reliability, 25 stories were chosen at random from the four sub-samples. The judges had rated all of these on the Conflict and Emphasis Scales, but only nineteen stories were rated on the Punishment Scale, and only thirteen on the Dominance Scale. (Stories were not rated when a scale was considered inapplicable or there was insufficient information on which to base a rating.) To establish a basis for comparing reliabilities, however, additional stories were chosen until there were 25 sets of ratings for each scale. Table 2 gives correlations between Judges A, B, and C.

Table 2

Interjudge Correlations on Rating Scales

(N=25 for each scale)

| Judges | Conflict Scale | Emphasis Scale | Punishment Scale | Dominance Scale |
|---|---|---|---|---|
| A with B | .66 | .71 | .61 | .72 |
| A with C | .68 | .71 | .43 | .77 |
| B with C | .85 | .74 | .91 | .67 |
| Average | .73 | .72 | .65 | .72 |

Notice that interjudge agreements vary on different scales. Judge A's ratings tended to deviate from those of Judges B and C, particularly on the Punishment Scale. (This shows the desirability of selecting the most reliable from a pool of judges, and/or extending training to increase reliability.)

Subsequent analysis was restricted to the Conflict and Emphasis Scales, because they (a) were applicable to virtually all stories, and (b) had consistently high reliability. A multi-factorial analysis of variance was applied to the ratings, accounting for the factors of location, sex of child, religion of child, and socio-economic status of parents. Because the subclass frequencies were not proportional, a special non-orthogonal procedure was used.

Results will be presented separately for each analytic method. In so doing, we shall specify (a) type of statistics employed, (b) kind of information obtained, and (c) the specific findings with the method. After this, the two sets of results will be compared as to consistencies and discrepancies.

## Categorizing Method

Analyzing the story contents into categories produces data at the nominal level of measure, necessitating non-parametric statistics (chi-square in this case). This allowed for only tests of differences among the sub-samples, without accounting for the factors of sex, religion, and socio-economic status.[7] Any confounding of these could not be determined.

Rather than combine the two German sub-samples, they were kept separated in each chi-square calculation. Combination would have been justified only if no significant differences appeared between the two sub-samples for the category in question. On the contrary, differences did appear between them on several categories.

Chi-squares were computed to provide the most meaningful information. Sometimes they tested simple presence or absence of a story quality, while other tests examined differences as to type of quality manifested in a story (e.g. presence or absence of

---

[7] Non-parametric analyses for several variables seem not to have been clearly described in the statistical literature. Components of chi-square analysis has been described by Wilson (1956), but this requires orthogonality. One possible way to determine effects of several variables is to break down the analysis into several separate comparisons. In each comparison, all independent variables are held constant except the one being examined (e.g. sex), and a chi-square is computed. The difficulties of this are (a) several comparisons must be made for each variable, which means general effects and interactions can only be inferred, and (b) very large samples are required, because the holding constant of variables necessitates the use of very reduced portions of the original samples. This method already has been applied to the Anderson data (Terhune, 1959a, 1959b, 1960).

punishment, as opposed to type of punishment administered). In several instances, low frequencies necessitated combining data from several items within a category.

Of the seventeen categories in the coding scheme, significant differences were found in only seven of them (see tables in Appendix H). Among these, however, are some interesting indications. To give an integrated picture of these results, we will first discuss the story characteristics which were common to all the subsamples. With this for a frame of reference, the unique features of stories from the respective locations will then be presented.

## Story Aspects Common to All Locations

In all the locations, the child authors generally interpreted the situation of the damaged axe as the fault of the younger brother, Billy (H-1).[8] Only about a tenth of the stories specified that Herbert was at fault, while a fifth indicated that really no one was to blame. When Herbert discovers his chipped axe, most stories indicated that he assumes Billy to be guilty, and his actions follow accordingly (H-2). If Herbert makes no such assumption and asks Billy about the axe, Billy most often responds truthfully (H-3).

Few stories in any location indicated that Herbert considers Billy's youth to be a factor in his damaging the axe (H-6). If Herbert does anything with Billy, he usually punishes him directly, or reports him to the parents (H-4). About half the stories involved punishment of Billy by either Herbert or the parents (H-8). Physical punishment was by far the most common in all locations, while verbal punishment was next preferred (H-9). Deprivation was uncommon in the stories, while a multiple punishment occurred rarely. In the few stories that stated Billy's reactions, he was usually described as cooperative, in that he apologizes, promises reform, or makes amends (H-10).

---

[8] Parentheses refer to tables in Appendix H.

Seldom does Herbert do anything about restoring the axe, either by himself or in cooperation with Billy or the parents (H-5). The latter were brought into less than half the stories, and they usually enter the situation through being asked by Herbert or Billy (H-11). Once in the story, the parents most often restore the axe, and to a lesser degree they advise the boys and reject Billy's behavior. Occasionally they provide information and reject Herbert's behavior (H-12).

## Aspects Unique to the Locations

While the stories of all locations exhibited the above elements in common, there were several important aspects of the stories on which the sub-samples were heterogeneous. From the seven categories in which the sub-samples differed significantly, the following story characteristics were derived for the four locations.

Drammen. From the Drammen stories emerges a strong image of Herbert, in which Herbert is self-sufficient more often than in the other sub-samples. His feelings tend to be more fully described, and he is definitely angry about the axe situation (H-13). If Billy is to be punished, Herbert is more likely to administer the punishment than in the Hamburg and Benton Harbor sub-samples (H-8). Yet, we also find that Herbert advises and pardons Billy more than in the other locations (H-4), and only Drammen stories mention comradeship between the brothers (in 7% of Drammen stories).[9] Regarding the axe, Drammen stories are among those most likely to describe Herbert as himself restoring the axe (H-5). The parents are brought into the stories less than elsewhere (H-11).

Benton Harbor. The characterization of Herbert in the Benton Harbor stories contrasts markedly with the Drammen results. Herbert tends to rely considerably more upon the parents, who assume

---

[9] Comradeship was examined in Category 36; data are not presented for this category because of insufficient frequencies.

a major role in many stories. While Herbert's feelings are speci-
fied more often than in the other sub-samples, they are often
feelings of sadness rather than anger (H-13). In dealing with his
younger brother, Herbert informs on him to the parents consider-
ably more (H-4), and the parents administer the punishment more
often (H-9) than in the other locations.

Regarding the axe, Herbert less frequently restores it
himself than in the stories from Drammen and Hamburg (H-5). We
also find a markedly greater tendency for the axe to be replaced
rather than repaired (H-15).

Munich. Munich stories emphasize Herbert's punishing
Billy (H-4). In fact, punishment of Billy is almost always admin-
istered by Herbert (H-8). As in Drammen, the parents are not often
brought into the story (H-11).

Herbert's feelings are generally not mentioned in the
Munich stories (H-13). Also, less attention is given to the axe;
as in the Benton Harbor stories, Herbert seldom restores it (H-5),
and the axe is generally restored less often than in any of the
other sub-samples (H-15).

Hamburg. The Hamburg sub-sample shares some story traits
in common with each of the other places. As in Munich, there is
a tendency for Herbert to behave in a hostile manner toward Billy.
He punishes or informs on Billy in 94% of the Hamburg stories (H-4),
and relatively more of them impute to Herbert unfulfilled inten-
tions of punishing Billy (H-7).

In Hamburg, as in Benton Harbor, the parents are often
involved (H-11), and Billy is punished by them (H-8). But the
Hamburg stories are also similar to those of Drammen, in the ex-
tent to which the axe is restored by Herbert (H-5), or indeed by
anyone at all (H-15).

## Rating Scale Method

### Preliminary Analysis

Prior to analyzing multifactorially, a simple analysis

of variance was made across locations (a) to enable direct comparison with the content category results, and (b) to clarify the advantages gained by multifactorial analysis. As in the content category analysis, the two German sub-samples were kept separate.

Table 3 presents sub-sample means and the results of the simple analysis of variance on the Conflict and Emphasis scales. Differences were significant only for the Emphasis scale, although they approached significance on the Conflict scale. Inspection of the means reveals that stories in the sub-samples tended to attribute conflict behavior to Herbert, and emphasize the axe rather than interpersonal aspects of the story. Specific indications are that the most conflict appeared in the Munich stories, while the least conflict was found in Drammen. Also, the Hamburg stories placed the most emphasis on the axe, and Munich the least.

## Multifactorial Analysis

The multifactorial analysis was made so as to determine the respective contributions of location, sex, religion, and socio-economic status. We remind the reader that by so doing, we attempt to (a) establish location differences free from confounding effects of the other variables, and (b) determine the relative importance of location compared with the other variables.

A somewhat reduced portion of the original sub-samples was used in the multifactorial analysis. As Table 1 showed, information on the respective factors was not known for all subjects, so only those subjects for which information was available were included.

Table 1 also showed that the two German sub-samples differed considerably in composition, especially on the religion variable. The Hamburg sub-sample was predominantly Protestant and the Munich mainly Catholic. This suggested combining them to obtain an adequately balanced German sub-sample. To justify this, t-tests were made between portions of each sub-sample, holding constant all factors save location. Frequencies were sufficient only for comparisons among Protestant girls at S.E.S.-2

Table 3

Simple Analysis of Variance on the Four Locations

a. <u>Conflict Scale</u> (score over 6 indicates tendency to conflict be-
havior by Herbert; under 6 indicates tendency to harmonious be-
havior)

|  | Mean | n |
|---|---|---|
| Drammen | 6.35 | 101 |
| Munich | 7.13 | 49 |
| Hamburg | 6.88 | 50 |
| Benton Harbor | 6.75 | 99 |

| Source of Variance | D.F. | S.S. | Mn. Sq. | F |
|---|---|---|---|---|
| Between locations | 3 | 23.17 | 7.72 | 2.41 (Not signif.) |
| Within locations | 295 | 946.49 | 3.20 | (.05 $<$ P $<$ .10) |
| Total | 298 | 969.66 | | |

b. <u>Emphasis Scale</u> (score over 3 indicates more emphasis on inter-
personal aspects; under 3 indicates more emphasis on axe)

|  | Mean | n |
|---|---|---|
| Drammen | 2.64 | 101 |
| Munich | 2.88 | 49 |
| Hamburg | 2.18 | 50 |
| Benton Harbor | 2.62 | 99 |

| Source of variance | D.F. | S.S | Mn. Sq. | P |
|---|---|---|---|---|
| Between locations | 3 | 13.00 | 4.33 | 4.20 P $<$ .01 |
| Within locations | 295 | 302.33 | 1.03 | |
| Total | 298 | 315.33 | | |

and among Protestant girls at S.E.S.-3. In neither instance were there significant differences between the Munich and Hamburg groups on either the Conflict or Emphasis scale.

Indices used. To indicate the contribution of each variable to the overall results, we statistically estimate the magnitude of the "effects". (The procedure is described in greater detail in chapter IV.) In so doing, we are estimating how much deviation from the "true" grand mean is associated with the specific levels of each variable. We assume first that the mean for any combination of factors may be found by adding to the grand mean the "effects" of the particular level of each variable. For example, the "true" mean of the group "American girl Protestants in the third socio-economic level" would be found by adding to the "grand mean" the effects respectively of (a) being an American, (b) being a girl, (c) being a Protestant, and (d) being in the third socio-economic level.

The assumption of simple additivity of main effects will not suffice if variables interact significantly, e.g., if differences between boys and girls were significantly greater in Germany than in the United States. Here we would have "interaction" between sex and location. To estimate the "true mean" of the American-girl-Protestants-S.E.S.-3 mentioned above, we would have then to add the "effect" of being an American girl.

The problem is thus to estimate each of the main and interaction effects, and to determine from these which variables are statistically significant. The magnitudes of the effects of the significant variables may then be compared to determine which variables influence the results more. In the presentation of rating scale results, then, we shall give both the estimates of effects and the analysis of variance based on the effects.

Conflict Scale. From the Conflict Scale results (table 4), we can make the following statements:

 (a) Sex was the main factor, while location was relatively less important. Religion was a significant but secondary factor. Socio-economic differences

Table 4

Results of Multifactorial Analysis of Conflict Scale Data

**a. Analysis of variance**

| Source | D.F. | S.S. | Mn. Sq. | F | Indication |
|---|---|---|---|---|---|
| Location | 2 | 26.44 | 13.22 | 4.02 | Signif. @ .05 |
| Sex | 1 | 26.80 | 26.80 | 8.17 | Signif. @ .005 |
| Religion | 2 | 16.12 | 16.12 | 4.91 | Signif. @ .05 |
| Socio-economic Status | 2 | 2.42 | 1.21 | 0.30 | N.S. |
| All main effects | 6 | 68.24* | | | |
| All interaction | 22 | 48.95 | 2.22 | 0.68 | N.S. |
| Between Cells | 28 | 117.19 | | | |
| Error | 249 | 817.01 | 3.28 | | |
| Total | 277 | 934.20 | | | |

\* Sum of S.S.'s for 4 factors will not equal S.S. for "all main effects" because of non-orthogonality.

**b. Estimates of effects** (more positive the effect, the more Herbert's behavior tends toward conflict)

| | | |
|---|---|---|
| Grand Mean | +6.885 | |
| Germans | +0.285 | |
| Americans | +0.132 | Range = .702 |
| Norwegians | -0.417 | |
| Boys | + .383 | |
| Girls | - .383 | Range = .766 |
| Catholics | + .156 | |
| Protestants | - .156 | Range = .312 |
| S.E.S.-1 | + .068 | |
| S.E.S.-2 | + .034 | (Not significant) |
| S.E.S.-3 | - .103 | |

were not significant.

(b) German stories had the most conflict, while the American ones had slightly less. Norwegian stories had considerably less conflict than either of the other locations.

(c) Boys' stories had more conflict than did the girls', and those by Catholics had more than those by Protestants.

(d) The additivity of effects was supported.

The significance of each variable was obtained from the analysis of variance, while ranges of the effects indicated relative importance. For example, the range of .702 between the German and Norwegian effects indicates the contribution of nationality. The German effect of +.285 indicates that German stories had the most conflict; adding the grand mean of 6.885 indicates a "true" German mean of 7.170.

Because interaction was not significant, the additivity model was supported. As an additional check, a multifactorial analysis of variance with orthogonal formulas was made. A sex x location interaction was indicated, but this was not supported in a special test by the fitting constants method (results not given here).

Emphasis Scale. Results for the Emphasis Scale (table 5) indicated no significant differences among locations; only socio-economic status and religion were significant variables. Again, no significant interactions were found. While the grand mean of 2.864 reveals that the stories in general tended to emphasize the axe rather than interpersonal relations, this tendency increased with socio-economic level (i.e., the higher socio-economic levels had stories with the most concern about the axe). Regarding religion, stories by Protestants emphasized the axe more than those by Catholics. Comparison of ranges shows that SES and religion are about equal in their contributions.

Comparisons with non-factorial results. We can now see the distinct advantage gained by the rating scale-multifactorial analysis. Not only have location differences been shown free from confounding effects of the other factors, but perspective has been

## Table 5

Results of Multifactorial Analysis of Emphasis Scale Data

a. **Analysis of variance**

| Source | D.F. | S.S | Mn. Sq. | F | Indication |
|---|---|---|---|---|---|
| Location | 2 | 3.58 | 1.79 | 1.53 | N.S. |
| Sex | 1 | 4.06 | 4.06 | 3.48 | N.S. |
| Religion | 1 | 5.11 | 5.11 | 4.38 | Signif. @ .05 |
| Socio-economic Status | 2 | 8.86 | 4.43 | 3.80 | Signif. @ .05 |
| All main effects | 6 | 14.92 | 2.49 | | |
| All interactions | 22 | 22.30 | 1.01 | 0.86 | N.S. |
| Between cells | 28 | 37.22 | | | |
| Error | 250 | 291.62 | 1.166 | | |
| Total | 278 | 328.84 | | | |

b. **Estimates of effects** (the more positive the effect, the more emphasis on interpersonal relations; negative effects indicate increased emphasis on the axe.)

| | | |
|---|---|---|
| Grand mean | +2.684 | |
| Norwegians | +0.113 | |
| Americans | +0.027 | N.S. |
| Germans | -0.140 | |
| Boys | +0.074 | N.S. |
| Girls | -0.074 | |
| Catholics | +0.198 | Range = .396 |
| Protestants | -0.198 | |
| S.E.S.-3 | +0.217 | |
| S.E.S.-2 | -0.013 | Range = .421 |
| S.E.S.-1 | -0.204 | |

gained by comparing the relative contributions of each. On the Conflict Scale, we now know that location is indeed a significant variable, but sex differences contribute at least as much. In contrast, location differences were not significant on the Emphasis Scale.

Inspection of the effects reveals that the three nationalities have nearly the same relation to each other as existed in the non-factorial analysis. On the Conflict Scale, for example, German stories have the most Conflict, the Americans a little less, and the Norwegians the least. This is surprising, in light of the fact that sex and religion were both important contributing factors, and both variables were considerably confounded in the nationality samples. However, scrutiny of the data reveals why this is so. From the sex and religion effects, we see that boys' stories had generally more conflict than did the girls', and Catholic stories had more conflict than the Protestant stories. Referring now to table 1, we see that the German sample had considerably more girls than did the other samples, which would tend to lower the German scores. But we also see that the German sample had more Catholics than did the others, which would tend to raise the scores. By such counterbalancing effects, which occurred on both scales, the relations among the nationality groups have apparently been little disturbed by the confounding. However, it takes little imagination to realize that had the disproportions on the relevant variables been in different directions, the confounding could have greatly distorted the apparent relations among the nationalities. Evidence is thus provided to demonstrate the necessity of accounting for disproportions on demographic factors in cross-national research.

## Comparison of Content Category and Rating Scale Results

Comparison between the two sets of data is somewhat difficult because in no case do a rating scale and content category assess exactly the same story quality. In addition, the former

are based on molar story configurations while the latter record specific details. Nevertheless, some interesting comparisons are possible, which we shall now present.

The Conflict Scale results are probably best compared with results from category 24 (Herbert's actions regarding Billy). From table H-4 in Appendix H, we may combine the "punish" and "tells on Billy" items as indicating stories with "conflict " tendencies. We can then order the four sub-samples on these percentages, and compare the percentages with sub-sample means on the Conflict Scale. These comparisons are made in table 6.

### Table 6

#### "Conflict" Data from Two Analytic Methods

| Sub-samples | Categ. 24 % "conflict" stories | Conflict Scale Mean | |
|---|---|---|---|
| | | Simple Analysis | Multifact. Analysis |
| Hamburg | 93.7% | 6.88 | ⎫ 7.17 |
| Munich | 86.2% | 7.13 | ⎭ |
| Benton Harbor | 86.4% | 6.75 | 7.02 |
| Drammen | 76.9% | 6.35 | 6.47 |

There seems to be a tendency for sub-samples with proportionately more conflict stories also to have a higher mean on the Conflict Scale, as we might expect. The multifactorial analysis of the Conflict Scale data later corroborated this ordering of locations, but more clearly showed it to be associated with location differences and not simply due to confounding of the other variables.

While no content category assesses relative emphasis in the stories, Category 37 (restoration of the axe) probably is the best indicator of concern with the axe. Therefore, the results for category 37 shall be compared with those of the Emphasis Scale.

From the data presented on Category 37 in table H-17, we may rank the four sub-samples according to percentages of stories in which the axe is restored. Likewise, from table 3 we obtain the means on the Emphasis Scale for the four sub-samples. These

data are presented in Table 7.

Table 7

"Emphasis " Data from Two Analytic Methods

| Sub-samples | Categ. 37 % stories in which axe restored | Emphasis Scale Mean Simple Analysis | Emphasis Scale Mean Multifact. Analysis |
|---|---|---|---|
| Hamburg | 52.0% | 2.18 | 2.54 (German) |
| Drammen | 41.6% | 2.64 | 2.80 |
| Benton Harbor | 35.4% | 2.62 | 2.71 |
| Munich | 26.6% | 2.88 | 2.54 (German) |

There appears a definite tendency for percentage of stories in which the axe is repaired to be negatively correlated with Emphasis Scale mean. As a lower Emphasis Scale score indicates more emphasis on things (the axe), these results are quite consistent. However, the multifactorial analysis showed story emphasis not to be associated with the location variable, so we might infer that the significant sub-sample differences on category 37 do not represent location differences, but probably result from confounding of the religion and socio-economic variables.

A probable effect of confounding may be noted for the two German sub-samples. While their Emphasis Scale means differed considerably, note in table 1 that they also had opposite distributions on the religious variable. The preponderance of Protestants in the Hamburg sub-sample would tend to decrease the Hamburg mean, while the dominance of Catholics in the Munich sub-sample would spuriously elevate the Munich mean. The gross difference between Emphasis Scale means for the two sub-samples is probably due mainly to these effects.

## Interpreting Results:  Limitations

While the emphasis of this thesis is on methodology, it seems necessary to comment on two important issues raised by the analytic results.  One concerns the imputing of causality, while

the other pertains to the basic objective of determining nationality differences.[10]

## Causality

It is imperative that the analyses reported in this thesis be recognized as of an associative or correlative nature. That is, we have established relations between certain demographic variables and written story characteristics. By no means, however, have these results proved any causal relations. To do so, we should show that manipulation of one variable is always followed by a certain consequence, but this is an operation virtually impossible in a cross-national field investigation. But the laboratory offers no ready solution either, for under what conditions could a lifetime of differential experiences due to sex, religion, and socioeconomic status be manipulated?

The point may seem trivial, then, that we shall probably never be able to establish causality regarding effects of demographic factors. In the absence of this, however, it seems necessary to rely on ecological generality as indication of the validity of our relations. That is, relations as found in this study must be checked over numerous other nationalities not included here. By this means, we may be more certain of the generality of the findings, and would be in a safer position for any inference we care to make about causality.

## Nationality Differences

One aim of cross-national studies is to discover the contribution of nationality or culture to differences of personality, attitudes, etc.. So also in this study, we attempted to clarify the effect of nationality per se by culling out the effects of other variables. The residual we loosely labeled as "location"

---

differences, referring to effects of being an American, a German, or a Norwegian. However, the question that is raised is this: if we have discovered that the factors examined contributed significantly to the variance among sub-samples, how can we be certain that other factors, which we did not examine, might not explain the remaining sub-sample variance? That is, the "location" differences obtained in the multifactorial analysis might be due to disproportions on other unknown variables, such as age, intelligence, and so on. This is a problem that has been of concern to other researchers in the cross-national area. Klineberg (1950) points out that there are those who deny that national differences in personality would exist were all relevant variables to be accounted for. The opinion expressed by Klineberg is that even after all relevant variables were controlled, nationality differences would still be found. However, due to the virtually infinite number of factors that could be examined in cross-national research, an absolute answer to this question will probably never be established.

What, then, can the social scientist do about the problem? It seems to this writer that cross-national studies can still be of value by attempting to account for as many variables theoretically seeming to be of importance. It seems quite beneficial, as in this study, to demonstrate that some "nationality" differences are associated with varying proportions on certain major variables. We also benefit in knowing that certain variables are at least as important as nationality factors. This study has established that sex, religion, and socio-economic factors ought to be accounted for in cross-national research, and subsequent studies may reveal other factors to be of importance also. The more such factors that are accounted for, the better we shall be able to judge whether the remaining sample differences are truly nationality differences.

It should be added that separation of nationality from other variables is feasible mainly within culturally-similar blocs, in which the same variables apply from nation to nation. That is, comparisons across grossly different cultures are likely to involve

interactions so complete between nationality and other factors that separation of main effects becomes meaningless. In comparing the United States and India, for example, separation of religion from nationality would not be possible, except for the trivial minorities who belong to religions common to both countries.

<div align="center">*         *         *</div>

The next chapter presents a detailed discussion of the multifactorial analysis used on the rating scale data. Presuming a working knowledge of the standard (orthogonal) analysis of variance, it is provided for the reader who wishes to understand and use the non-orthogonal method. Those who are uninterested in the technicalities may skip over to chapter V.

## IV. MULTIFACTORIAL ANALYSIS FOR
## DISPROPORTIONATE DATA

An important consideration in cross-national research is the confounding of demographic variables. Proportions on relevant factors often vary across different national populations, and representative sampling may be expected to reflect these variations. Consequently, the confounding of nationality differences with differences on other factors may lead to misinterpretation of results if the disproportionality is not taken into account. This problem is also likely in other field studies, for relevant variables cannot always be controlled. Statistical methods for handling the problem, although available for some time, are apparently little known among psychologists.[11] In this chapter, therefore, a method is presented which enables analysis of variance to be used with non-orthogonal, i.e., disproportional data. Called the Method of Fitting Constants, it seems to be the most generally applicable of the non-orthogonal methods. The technique will be explained using the Conflict Scale data as an example, and will be based on Bennett and Franklin (1954), Scheffe (1959), Stevens (1948), and Yates (1934).

Advantages of the Method

Valuable features of the Method of Fitting Constants are:

(a) It provides estimates of main effects, enabling comparisons of magnitude.

(b) It provides tests for interactions.

(c) It can be used even when some cells (combinations of factors) have zero frequency.

(d) It can be readily used with any number of factors.

---

[11] Statistics textbooks commonly discuss the Method of Weighted Means, while not always so designating it. Although simple, it becomes less precise as the sub-class numbers become more unequal (Yates, 1934, p. 57), and it cannot be used if there are any empty cells (Scheffe, 1959, p. 362).

(e) With the advent of electronic computers, it is cap-
able of easy solution.

## General Approach

The Method of Fitting Constants begins with the addi-
tivity model, which is based on the hypothesis that the true mean
for any combination of factors may be obtained by adding to the
grand mean the "effects" of each contributing factor. If this is
true, each variable is independent, i.e., there is no interaction
(Scheffé, 1959).

Under the additivity hypothesis, the effects are esti-
mated and significance tests made. If interaction is found non-
significant, the additivity model is considered tenable for the
significant variables. If, however, interaction is found signifi-
cant, the additivity model is rejected, and the data must be ex-
amined to determine which variables significantly interact. These
steps will now be demonstrated on the Conflict Scale data.

## The Additivity Model

The model is written for the mean of any combination of
location, sex, religion, and socio-economic status, which are the
four variables in the problem. Symbolically, the model is written,

$$\mu_{ijkl} = \mu + E_{L_i} + E_{S_j} + E_{R_k} + E_{SES_l}$$

where $\mu_{ijkl}$ = the Conflict Scale Mean for a particular combination
of location i, sex j, religion k, and socio-economic
status l. As there are three locations, two sexes,
two religions, and three SES's, there theoretically
can be 36 such means.

$\mu$ = true grand mean on the Conflict Scale

$E_{L_i}$ = effect of a particular location (i = 1,2,3)

$E_{S_j}$ = effect of a particular sex (j = 1,2)

$E_{R_k}$ = effect of a particular religion (k = 1,2)

$E_{SES_l}$ = effect of a particular socio-economic status
(l = 1,2,3)

For estimating the values of the grand mean and each of the effects, the above equation is rewritten as

$$M_{ijkl} = M + a_i + b_j + c_k + d_l$$

In this equation, each term is the sample estimate of its counterpart in the first equation, which specified only population parameters.

## Solving for Effects

Having assumed the additivity model, the next step is to solve for the unknown constants $M$, $a_i$, $b_j$, $c_k$, and $d_l$. By the method of maximum likelihood, the estimates are obtained by minimizing the variance around $\mu_{ijkl}$. If we let $X_{ijklm}$ symbolize any obtained value within cell (combination) i-j-k-l, the best estimates of the unknowns are obtained by minimizing

$$\Sigma_{ijklm} (X_{ijklm} - \mu - E_{L_i} - E_{S_j} - E_{R_k} - E_{SES_l})^2$$

The summation sign indicates that the parenthesized values are added over all individual values in all cells.

By methods of calculus (Bennett & Franklin, 1954; Scheffé, 1959), it can be shown that the above sum of squared deviations will be minimized when the estimated mean and effects will solve the following set of equations.

(1) $NM + n_{1...}a_1 + n_{2...}a_2 + n_{3...}a_3 + n_{.1..}b_1 + n_{.2..}b_2$
$+ n_{..1.}c_1 + n_{..2.}c_2 + n_{...1}d_1 + n_{...2}d_2 + n_{...3}d_3 \qquad = \Sigma_{ijklm}X$

(2) $n_{1...}M + n_{1...}a_1 + (0)a_2 + (0)a_3 + n_{11..}b_1 + n_{12..}b_2$
$+ n_{1.1.}c_1 + n_{1.2.}c_2 + n_{1..1}d_1 + n_{1..2}d_2 + n_{1..3}d_3 \qquad = \Sigma X_1....$

(3) $n_{2...}M + (0)a_1 + n_{2...}a_2 + (0)a_3 + n_{21..}b_1 + n_{22..}b_2$
$+ n_{2.1.}c_1 + n_{2.2.}c_2 + n_{2..1}d_1 + n_{2..2}d_2 + n_{2..3}d_3 \qquad = \Sigma X_2....$

(4) $n_{3...}M + (0)a_1 + (0)a_2 + n_{3...}a_3 + n_{31..}b_1 + ... + n_{3...}d_3 = X_3....$

(5) $n_{.1..}M + n_{11..}a_1 + n_{21..}a_2 + n_{31..}a_3 + n_{.1..}b_1$
$+ (0)b_2 + \qquad . \qquad . \qquad . \qquad + n_{.1.3}d_3 \qquad = X_{.1}...$

$$. \qquad . \qquad . \qquad . \qquad . \qquad .$$

(11) $n_{...3}M + n_{1..3}a_1 + n_{2..3}a_2 + n_{3..3}a_3 + ... + n_{...3}d_3 = \Sigma X_{...3}.$

This imposing set of equations need not horrify the reader, for the complete set may be derived by a simple and logical procedure. A table is made as in table 8, in which the equations estimating each cell mean are written. The marginal totals for each variable are also listed. With the table, the set of equations can be written, based on the fact that each marginal total is the sum of all the individual cell totals comprising that group. For example, the total $\sum X$ for the U.S.A. group is the sum of all the individual cell totals in the U.S.A. sub-sample. Symbolically,

$$\sum X_{\text{U.S.A.}} = 558.3 = \sum X_{\text{USA-boys-Prot-SES1}} + \sum X_{\text{USA-girls-Prot-SES1}}$$
$$+ \sum X_{\text{USA-boys-Cath-SES1}} + \sum X_{\text{USA-girls-Cath-SES1}}$$
$$+ \sum X_{\text{USA-boys-Prot-SES2}} + \sum X_{\text{USA-girls-Prot-SES2}}$$
$$+ \ldots$$
$$+ \sum X_{\text{USA-boys-Cath-SES3}} + \sum X_{\text{USA-girls-Cath-SES3}}$$

As $\sum X = n \bar{X}$, we can substitute for the $\sum X$ of each cell its theoretical equivalent $n_{ijkl} \times \mu_{ijkl}$. As $\mu_{ijkl}$ is estimated by $(M + a_i + b_j + c_k + d_l)$, we may write

$$\sum X_{\text{USA}} = 3(M + a_1 + b_1 + c_1 + d_1) + 2(M + a_1 + b_2 + c_1 + d_1)$$
$$+ 4(M + a_1 + b_2 + c_2 + d_1) + \ldots + 3(M + a_1 + b_2 + c_2 + d_3)$$
$$= 558.3$$

This equation, when the coefficients of each constants are summed, reduces to

$$\sum X_{\text{USA}} = 81M + 81a_1 + (0)a_2 + (0)a_3 + 38b_1 + 43b_2 + 57c_1 + 24c_2$$
$$+ 11d_1 + 25d_2 + 45d_3 = 558.3$$

Notice the simplicity of the coefficients. That is, the coefficient

| of M is the total no. of Americans in the U.S.A. group |
| " $a_1$ " " " " " " " " " " |
| " $a_2$ " " " " " Norwegians " " " " (none) |
| " $a_3$ " " " " " Germans " " " " (none) |
| " $b_1$ " " " " " boys " " " " |
| " $b_2$ " " " " " girls " " " " |
| " $c_1$ " " " " " Protestants " " " " |
| " $c_2$ " " " " " Catholics " " " " |

## Table 8: Data Layout for Solving Main Effects on Conflict Scale

a. Estimates of cell means, and cell frequencies

| | | U.S.A. Boys | U.S.A. Girls | Norway Boys | Norway Girls | Germany Boys | Germany Girls |
|---|---|---|---|---|---|---|---|
| SES 1 | Prot | $M+a_1+b_1+c_1+d_1$ $n=3$ | $M+a_1+b_2+c_1+d_1$ $n=2$ | $M+a_2+b_1+c_1+d_1$ $n=10$ | $M+a_2+b_2+c_1+d_1$ $n=8$ | $M+a_3+b_1+c_1+d_1$ $n=4$ | $M+a_3+b_2+c_1+d_1$ $n=6$ |
| | Cath | $M+a_1+b_1+c_2+d_1$ $n=4$ | $M+a_1+b_2+c_2+d_1$ $n=2$ | $M+a_2+b_1+c_2+d_1$ $n=0$ | $M+a_2+b_2+c_2+d_1$ $n=0$ | $M+a_3+b_1+c_2+d_1$ $n=0$ | $M+a_3+b_2+c_2+d_1$ $n=6$ |
| SES 2 | Prot | $M+a_1+b_1+c_1+d_2$ $n=6$ | $M+a_1+b_2+c_1+d_2$ $n=10$ | $M+a_2+b_1+c_1+d_2$ $n=21$ | $M+a_2+b_2+c_1+d_2$ $n=16$ | $M+a_3+b_1+c_1+d_2$ $n=7$ | $M+a_3+b_2+c_1+d_2$ $n=15$ |
| | Cath | $M+a_1+b_1+c_2+d_2$ $n=4$ | $M+a_1+b_2+c_2+d_2$ $n=5$ | $M+a_2+b_1+c_2+d_2$ $n=0$ | $M+a_2+b_2+c_2+d_2$ $n=0$ | $M+a_3+b_1+c_2+d_2$ $n=6$ | $M+a_3+b_2+c_2+d_2$ $n=15$ |
| SES 3 | Prot | $M+a_1+b_1+c_1+d_3$ $n=15$ | $M+a_1+b_2+c_1+d_3$ $n=21$ | $M+a_2+b_1+c_1+d_3$ $n=27$ | $M+a_2+b_2+c_1+d_3$ $n=19$ | $M+a_3+b_1+c_1+d_3$ $n=12$ | $M+a_3+b_2+c_1+d_3$ $n=8$ |
| | Cath | $M+a_1+b_1+c_2+d_3$ $n=6$ | $M+a_1+b_2+c_2+d_3$ $n=3$ | $M+a_2+b_1+c_2+d_3$ $n=0$ | $M+a_2+b_2+c_2+d_3$ $n=0$ | $M+a_3+b_1+c_2+d_3$ $n=4$ | $M+a_3+b_2+c_2+d_3$ $n=13$ |

b. Marginal totals

| | | | |
|---|---|---|---|
| U.S.A. : | $\Sigma X = 558.3$ | Protestants : | $\Sigma X = 1387.6$ |
| Norway : | $\Sigma X = 641.0$ | Catholics : | $\Sigma X = 486.0$ |
| Germany: | $\Sigma X = 674.3$ | SES1 : | $\Sigma X = 307.0$ |
| Boys : | $\Sigma X = 911.5$ | SES2 : | $\Sigma X = 713.1$ |
| Girls : | $\Sigma X = 962.1$ | SES3 : | $\Sigma X = 853.5$ |

$N = 278$    $\Sigma_{ijklm}X = 1873.6$

of $d_1$ is the total no. of SES1's in the U.S.A. group
" $d_2$ " " " " " " SES2's " " " "
" $d_3$ " " " " " " SES3's " " " "

Now, similar equations may be written for $\Sigma X_{Norway}$, $\Sigma X_{Germany}$, $\Sigma X_{boys}$, $\Sigma X_{girls}$, $\Sigma X_{Prot}$, etc., and also for the grand total, $\Sigma_{ijklm} X$. These equations are not, however sufficient to solve for the various effects. To obtain the remaining equations necessary, the assumption is made that the sum of each set of effects is zero. That is,

$$a_1 + a_2 + a_3 = 0$$

$$b_1 + b_2 = 0$$

$$c_1 + c_2 = 0$$

$$d_1 + d_2 + d_3 = 0$$

The full set of equations is presented in table 9. For a check on correctness, the respective sums of the location, sex, religion, and S.E.S. equations should each equal the grand total equation. This fact also indicates that the sets of equations are not independent, i.e., they provide overlapping or redundant information. To solve for the unknowns, therefore, we need only the following:

(a) any two of the three location equations
(b) either of the two sex equations
(c) either of the two religion equations
(d) any two of the three S.E.S. equations
(e) the grand total equation
(f) all side condition equations

Thus, we can solve the problem using equations 1, 2, 4, 6, 8, 9, 11, 12, 13, 14, and 15. (Always, as many independent equations are needed as there are unknowns.)

Solving these equations by standard desk calculator methods is a tedious task, and even a simplified iterative method presented by Stevens (1948) is still time consuming. (About 30 hours were required for the above equations.) Use of an electronic computer is the most expedient method, as a data tape for the equations may be made in about fifteen minutes. The computer will then solve the

## Table 9

### Equations for Estimating Grand Mean and Effects: Conflict Scale

**Location equations:**

$$(1)\quad 81M + 81a_1 + (0)a_2 + (0)a_3 + 38b_1 + 43b_2 + 57c_1 + 24c_2 + 11d_1 + 25d_2 + 45d_3 = 558.3$$

$$(2)\quad 101M + (0)a_1 + 101a_2 + (0)a_3 + 58b_1 + 43b_2 + 101c_1 + (0)c_2 + 18d_1 + 37d_2 + 46d_3 = 641.0$$

$$(3)\quad 96M + (0)a_1 + (0)a_2 + 96a_3 + 33b_1 + 63b_2 + 52c_1 + 44c_2 + 16d_1 + 43d_2 + 37d_3 = 674.3$$

**Sex equations:**

$$(4)\quad 129M + 38a_1 + 58a_2 + 33a_3 + 129b_1 + (0)b_2 + 105c_1 + 24c_2 + 21d_1 + 44d_2 + 64d_3 = 911.5$$

$$(5)\quad 149M + 43a_1 + 43a_2 + 63a_3 + (0)b_1 + 149b_2 + 105c_1 + 44c_2 + 24d_1 + 61d_2 + 64d_3 = 962.1$$

**Religion equations:**

$$(6)\quad 210M + 57a_1 + 101a_2 + 52a_3 + 105b_1 + 105b_2 + 210c_1 + (0)c_2 + 33d_1 + 75d_2 + 102d_3 = 1387.6$$

$$(7)\quad 68M + 24a_1 + (0)a_2 + 44a_3 + 24b_1 + 44b_2 + (0)c_1 + 68c_2 + 12d_1 + 30d_2 + 26d_3 = 486.0$$

**S.E.S. equations:**

$$(8)\quad 45M + 11a_1 + 18a_2 + 16a_3 + 21b_1 + 24b_2 + 33c_1 + 12c_2 + 45d_1 + (0)d_2 + (0)d_3 = 307.0$$

$$(9)\quad 105M + 25a_1 + 37a_2 + 43a_3 + 44b_1 + 61b_2 + 75c_1 + 30c_2 + (0)d_1 + 105d_2 + (0)d_3 = 713.1$$

$$(10)\quad 128M + 45a_1 + 46a_2 + 37a_3 + 64b_1 + 64b_2 + 102c_1 + 26c_2 + (0)d_1 + (0)d_2 + 128d_3 = 853.5$$

**Grand total equation:**

$$(11)\quad 278M + 81a_1 + 101a_2 + 96a_3 + 129b_1 + 149b_2 + 210c_1 + 68c_2 + 45d_1 + 105d_2 + 128d_3 = 1873.6$$

**Side condition equations:**

$$(12)\quad 0 + a_1 + a_2 + a_3 + 0 + 0 + 0 + 0 + 0 + 0 + 0 = 0$$

$$(13)\quad 0 + 0 + 0 + 0 + b_1 + b_2 + 0 + 0 + 0 + 0 + 0 = 0$$

$$(14)\quad 0 + 0 + 0 + 0 + 0 + 0 + c_1 + c_2 + 0 + 0 + 0 = 0$$

$$(15)\quad 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + d_1 + d_2 + d_3 = 0$$

$M$ = grand mean
$a_1$ = American effect
$a_2$ = Norwegian effect
$a_3$ = German effect

$b_1$ = boy effect
$b_2$ = girl effect
$c_1$ = Protestant effect
$c_2$ = Catholic effect

$d_1$ = S.E.S. 1 effect
$d_2$ = S.E.S. 2 effect
$d_3$ = S.E.S. 3 effect

problem quickly and exactly.    The effects presented in table 10 were so obtained.


## Testing for Significance

Analysis of variance is used for significance tests.  To determine the sum of squares for each factor, it is necessary first to obtain the sum of squares for all main effects combined.  This is given by the following formula (Yates, 1934):

$$S.S._{\text{Main Effects}} = M_{ijkl\text{m}}X + a_1\Sigma X_{USA} + a_2\Sigma X_{Norway} + a_3\Sigma X_{Germany} + b_1\Sigma X_{boys} + \cdots + d_3\Sigma X_{SES3} - \frac{(\Sigma X_{ijklm})^2}{N}$$

Substituting the values from tables 8 and 10,

$$S.S._{\text{Main Effects}} = 6.885(1873.6) + .132(558.3) - .417(641.0)$$
$$+ .285(674.3) + .383(199.5) + \cdots$$
$$- .103(853.5) - \frac{(1873.6)^2}{278} = 68.24$$

Now, to derive the SS's for each factor, sums of squares based on the marginal totals are first calculated.  Formulas for standard (orthogonal) analysis of variance are used.  Thus, for the sex variable,

$$SS_{\substack{Sex,\\ \text{Marginal}}} = \frac{(\Sigma X_{boys})^2}{n_{boys}} + \frac{(\Sigma X_{girls})^2}{n_{girls}} - \frac{(\Sigma X_{ijklm})^2}{N}$$
$$= \frac{(911.5)^2}{129} + \frac{(962.1)^2}{149} - \frac{(1873.6)^2}{278} = 25.62$$

The marginal-based SS's for the other variables are calculated similarly.  These are confounded because the data are non-orthogonal, but the unconfounded sum of squares for each variable may be obtained as follows:

$$S.S._{Location} = SS_{\substack{\text{Main}\\ \text{Effects}}} - SS_{\substack{Sex,\\ \text{Marg.}}} - SS_{\substack{Rel.\\ \text{Marg.}}} - SS_{\substack{SES\\ \text{Marg.}}}$$

$$S.S._{Sex} = SS_{\substack{\text{Main}\\ \text{Effects}}} - SS_{\substack{Loc.\\ \text{Marg.}}} - SS_{\substack{Rel.\\ \text{Marg.}}} - SS_{\substack{SES\\ \text{Marg.}}}$$

$$S.S._{Religion} = SS_{\substack{\text{Main}\\ \text{Effects}}} - SS_{\substack{Loc.\\ \text{Marg.}}} - SS_{\substack{Sex,\\ \text{Marg.}}} - SS_{\substack{SES\\ \text{Marg.}}}$$

Table 10

Results of Multifactorial Analysis of Conflict Scale Data

(These data are repeated from table 4 for easy reference)

a. Estimates of effects (more positive the effect, the more
Herbert's behavior tends toward conflict)

| | |
|---|---|
| Grand Mean | +6.885 |
| Germans | +0.285 |
| Americans | +0.132 |
| Norwegians | -0.417 |
| Boys | +0.383 |
| Girls | -0.383 |
| Catholics | +0.156 |
| Protestants | -0.156 |
| S.E.S.-1 | +0.068 |
| S.E.S.-2 | +0.034 |
| S.E.S.-3 | -0.103 |

Note: These effects were obtained by using program L7-S on
Michigan State University's MISTIC computer.

b. Analysis of variance

| Source | D.F. | S.S. | Mn. Sq. | F | Indication |
|---|---|---|---|---|---|
| Location | 2 | 26.44 | 13.22 | 4.03 | Signif. .05 |
| Sex | 1 | 26.80 | 26.80 | 8.17 | Signif. .005 |
| Religion | 1 | 16.12 | 16.12 | 4.91 | Signif. .05 |
| SES | 2 | 2.42 | 1.21 | 0.30 | N.S. |
| All main effects | 6 | 68.24* | | | |
| All interaction | 22 | 48.95 | 2.22 | 0.68 | N.S. |
| Between cells | 28 | 117.19 | | | |
| Error | 249 | 817.01 | 3.28 | | |
| Total | 277 | | | | |

*Sum of S.S.'s for 4 factors will not equal S.S. for "all main
effects" because of non-orthogonality.

$$S.S._{SES} = SS_{\substack{Main \\ Effects}} - SS_{\substack{Loc. \\ Marg.}} - SS_{\substack{Sex, \\ Marg.}} - SS_{\substack{Rel., \\ Marg.}}$$

With these sums of squares, the analysis of variance table may then be written, as in table 10. In contrast to analysis with orthogonal data, the four sums of squares from the main factors will not add to the sum of squares for "all main effects".

The between, error, and total sums of squares are calculated in the usual way, for they are not affected by non-orthogonality. (It is to be noted that the obtained cell means are unbiased and efficient estimates of the true cell means $\mu_{ijkl}.$) The degrees of freedom for between cells are one less than the number of cells containing data, or

$$(i \times j \times k \times l) - (\text{no. zero-frequency cells}) - 1$$

The error degrees of freedom are the difference between the total D.F. and the "between " D.F., or $\Sigma_{ijkl}(n_{ijkl} - 1)$. Tests of significance are made in the usual way, using the error mean square in the denominator of the F-ratio for every factor.

## Testing Interaction

An overall interaction sum of squares is found by subtracting the S.S. for "all main effects" from the "between" S.S. The degrees of freedom are obtained similarly. Significance of interaction is then tested by an F-ratio, using the error mean square in the denominator. Table 10 indicates no interaction in the Conflict Scale data.

If the F-ratio had been significant or near-significant, further tests would have been required to determine those factors with significant interaction. (Near-significant results indicate that effects of one or more specific interactions may have been diluted by pooling with other non-significant interactions.)

Estimation and significance tests for interaction among specific combinations of factors are similar to the tests for main effects. Instead of using the marginal totals, however, equations for the interaction effects involve various sub-totals among combinations of factors.

It should be possible to write sets of equations to solve
for the various first- and second-order interactions (involving sets
of two variables and three variables respectively). However, the
number of interaction effects becomes so extensive in multi-factor
problems, that solution becomes quite complex, and the value of so
many data is questionable. For example, in the analysis of the
Conflict Scale data, solution could be made for 37 first-order and
60 second-order interaction effects. In lieu of such extensive
analysis, a simplified approach based on Steven's (1948) method is
proposed. This involves first making an analysis of variance which
ignores the non-orthogonality, using standard formulas. (Some of
the necessary calculations will already have been made in the main
effects solutions.) The results will be confounded, but will at
least indicate which interactions may be significant. These are
then examined more carefully.

To illustrate the procedure, suppose that in the Conflict
Scale Problem there had been indications that the sex x religion
interaction was significant. The model to be tested would then in-
clude this interaction effect. Omitting the non-significant SES
variable, the model is as follows:

$$\mu_{i(jk)} = \mu + a_i + (b_j c_k)$$

where $\mu_{i(jk)}$ = the true grand mean for a given combination of
location, sex, and religion

$\mu$ = true grand mean

$a_i$ = a specific location effect

$(b_j c_k)$ = the effect for a particular combination of sex and
religion. There are four such effects.

In the model, main effects of sex and religion have been
included in the interaction term to simplify the analysis. With
the model, a table of estimates for the cell means may be written
as in table 11. As in the main effects solution, the table is
used to write the necessary equations. For example, the equations
for Protestant boys would be as follows:

Table 11

Data Layout for Solving Sex x Religion Interaction
on Conflict Scale

a. Estimates of cell means, and cell frequencies

| | U.S.A. | | Norway | | Germany | |
|---|---|---|---|---|---|---|
| | Boys | Girls | Boys | Girls | Boys | Girls |
| Prot | $M+a_1+b_1c_1$ | $M+a_1+b_2c_1$ | $M+a_2+b_1c_1$ | $M+a_2+b_2c_1$ | $M+a_3+b_1c_1$ | $M+a_3+b_2c_1$ |
| | n = 24 | n = 33 | n = 58 | n = 43 | n = 23 | n = 29 |
| Cath | $M+a_1+b_1c_2$ | $M+a_1+b_2c_2$ | $M+a_2+b_1c_2$ | $M+a_2+b_2c_2$ | $M+a_3+b_1c_2$ | $M+a_3+b_2c_2$ |
| | n = 14 | n = 10 | n = 0 | n = 0 | n = 10 | n = 34 |

b. Marginal totals

U.S.A.:   $\Sigma X = 558.3$

Norway:   $\Sigma X = 641.0$

Germany:   $\Sigma X = 674.3$

Boy-Prot:   $\Sigma X = 731.6$

Girl-Prot:   $\Sigma X = 656.0$

Boy-Cath:   $\Sigma X = 179.9$

Girl-Cath:   $\Sigma X = 306.1$

$\Sigma X_{Prot\text{-}boys}$ = 731.6

= $24(M + a_1 + b_1c_1) + 58(M + a_2 + b_1c_1) + 23(M + a_3 + b_1c_1)$

= $105M + 24a_1 + 58a_2 + 23a_3 + 105b_1c_1$

The rest of the equations are similarly written, yielding four equations for sex x religion interaction, three location equations, and one grand total equation. We also add the side conditions:

$$a_1 + a_2 + a_3 = 0$$
$$b_1c_1 + b_1c_2 + b_2c_1 + b_2c_2 = 0$$

Omitting one of the interaction and one of the location equations, seven independent equations for solving the seven unknowns remain. (Note that new estimates of $a_i$ would be obtained.) Again, an electronic computer expedites the solution. With the estimated effects, the sum of squares for the location and sex x religion effects <u>combined</u> are obtained by the equation

$$
\begin{aligned}
S.S._{\substack{Loc, \\ SxR}} = \; & M\Sigma_{ijkl}X + a_1 \Sigma X_{USA} + a_2 \Sigma X_{Norway} + a_3 \Sigma X_{Germany} \\
& + (b_1c_1)\Sigma X_{Boy\text{-}Prot} + (b_1c_2)\Sigma X_{Boy\text{-}Cath} \\
& + (b_2c_1)\Sigma X_{Girl\text{-}Prot} + (b_2c_2)\Sigma X_{Girl\text{-}Cath} \\
& - \frac{(\Sigma X_{ijkl})^2}{N}
\end{aligned}
$$

To obtain the sum of squares for <u>just</u> the sex x religion interaction, from the combined location-SxR sum of squares (above) is subtracted the sum of squares for the location, sex, and religion main effects. The latter is the same as the $S.S._{Main\ Effects}$ computed earlier, except that the socio-economic status component is omitted. Thus:

$$S.S._{SxR} = S.S._{\substack{Loc, \\ SxR}} - S.S._{Main\ Effects}$$

The degrees of freedom for the sex x religion interaction is calculated in the usual way:

$$d.f._{SxR} = (j\text{-}1)(k\text{-}1) = (2\text{-}1)(2\text{-}1) = 1$$

where j = no. of sexes

k = no. of religions

To test the significance of the interaction, an F-ratio is obtained by dividing the sex x religion mean square by the error mean square, as usual.

# V. DISCUSSION

Having now presented the content category and rating scale methods of content analysis, and the comparative results, we will now review the findings from a methodological viewpoint. In so doing, we will attempt to answer each of the questions posed in Chapter I.

## 1. What kind of information can the two methods provide?

The information provided by the two analytic methods seems to vary in three ways: (a) in the factors that can be accounted for, (b) in the level of abstraction attainable, and (c) in the possible scope of application.

Factors accounted for. Perhaps the most outstanding contribution of the rating scale analysis was that it enabled us to determine the contribution of several factors additional to location in producing story differences. While the content category method could only reveal gross sub-sample differences, the analysis of variance on the rating scale data showed that sex was at least as important as location, and socio-economic and religious factors were also important. Such knowledge, by eliminating the confounding effects of certain extraneous variables, facilitates interpretation of results. In addition, we gain perspective on the importance of location by comparing the magnitude of its effect with those for the other variables.

Finally, because the rating analysis reveals that certain factors produce significant differences, it warns that the content category data must be interpreted with caution; regional differences across sub-samples may confound factors other than location.

Level of abstraction. Rating scales were found more applicable to assessment in terms of psychological concepts than were the content categories. While in theory it should be possible to apply content categories to the abstract psychological aspects of the stories, in practice such categories were difficult to code. They were usually eliminated from the coding scheme because of low

48

inter-judge reliability, and some were converted into rating scales. In order to raise their reliability, it would seem necessary to provide extensive instructions to the coder for interpreting specific details in terms of the psychological concept under consideration. Thus, the inference from observables to concept would be built into the coding scheme. The writer prefers to leave this inferring to the judge, who may rate on the basis of complete story context and not be confined by rigid adherence to a set of rules.

Scope of application. The scope of application that the two analytic methods can provide can be discussed in terms of (a) within-story application, and (b) across-story application.

(a) By "within-story application", we refer to the assessment of the various aspects of completions to a single incomplete story (story B-4 in this study). Here, content categegories were found to have a decided advantage over rating scales, for a content category could be readily devised for assessing almost any aspect of the story completions in which we were interested.[12] Furthermore, each category could be analyzed in different ways to provide different kinds of information (e.g., tables H-8, H-11, H-13 in Appendix H).

Rating scales, in contrast to content categories, were found applicable to very few aspects of the stories. Thus many of the proposed rating scales were not useable because their concepts were found inapplicable to the stories, or because the stories generally contained insufficient information upon which to base a rating.

(b) "Across-story application" refers to the ability to apply our assessment tool to different incomplete stories, desirable for determining consistency of response. On this matter we

---

[12] We realize that this can be overdone, for the researcher may get carried away in his zeal to record all story aspects. This is mentioned by Katz, who states that "The major error of the novice is his fear that he will lose some of the richness of his materials if he does not cover all phenotypical details. In general, the more elaborate and the more detailed the coding, the less useful it is in analysis." (Festinger and Katz, 1953, P. 90).

must infer cautiously, for this study did not actually involve cross-story applications. It would seem, however, that because rating scales may be used to assess somewhat more abstract concepts than do content categories, they would be more applicable to different incomplete stories. This would depend on the particular rating scale. The Conflict Scale, for example, could probably be applied across stories with only minor modifications, while it is unlikely that the Emphasis Scale could be so applied.

2. **Is there a difference in the time required by the two methods?**

The time consumed in developing, recording, and analyzing results per content category was found to be considerably less than that for rating scales. It is estimated that, from development of coding manual to production of final results, roughly 14 man-hours were spent per content category. In contrast, the equivalent procedures per rating scale required about 50-55 man-hours, using an electronic computer for the fitting of constants in the analysis of variance. This latter figure could probably be diminished if IBM equipment were used for some clerical operations (as was done in the content category analysis), but total time would probably still exceed that for the content category procedure.

In comparing time requirements of the two methods, content categories seem preferable, **if** we have no need to account for several variables, **if** we need not assess on abstract variables, etc.. In this particular study, however, the additional information gained from the rating scale analysis is considered by the writer to be worth the extra time.

3. **Are there differences in the inter-coder reliability of the two methods?**

With different indices of inter-judge reliability employed in the two methods, direct comparison is of course not possible. However, both methods appear to have relatively high reliability and to be about equally reliable. Agreement of 92% was obtainable with the content category method, while inter-judge correlations for the

rating scales averaged around .72. The latter figure, based on 25
ratings, would be equivalent to .91 for 100 ratings (Spearman-Brown
formula). Results indicated that reliability could be improved fur-
ther with proper training of judges.

4. What statistical analyses can be used with the two methods?

At this point we need not discuss further the application
of different statistics for the two methods. We will simply point
out once again that correlation analysis might also have been util-
ized with the rating scales, in which the contribution of the var-
ious factors would have been assessed via partial correlations.

5. Is one method superior to the other in detecting or showing dif-
ferences between different samples?

Ability to detect differences among samples is a function
of the statistical tests employed. As Siegel (1956) indicates,
there is no way to compare the power of the chi-square test with
other statistical tests. This point will not, therefore, be pur-
sued further.

Worth a brief comment here is the fact that several of
the content categories are not useful in showing sample differences
without having large samples. These are the categories which
utilized only a portion of the total sample, such as category 23
(communication by Billy). This category attempts to assess dif-
ferences in whether Billy tells the truth or not about the axe, and
is necessarily applicable only to those stories in which Billy com-
municates. Although large percentage differences were found on this
category (table H-3 in Appendix H), the n's were too small for dif-
ferences to be significant. Thus, this category has little dif-
ferential power without large sub-samples.

A point not always realized by the statistically unsoph-
isticated, although it may seem trivial to the statistically know-
ledgeable reader, is the fact that percentage differences, as were
used in results of the content category analysis, generally appear
more impressive than do mean differences, as were found with rating

scales. We can illustrate this by an example. Suppose that we had
two samples of stories scored on a five-point rating scale. Results
could be presented in terms of the percentage of stories receiving
different ratings, as shown below.

|  | Scale Point | | | | | Scale Mean |
| --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 |  |
| Sample A (N=100) | 10% | 25% | 45% | 10% | 10% | 2.85 |
| Sample B (N=100) | 10% | 45% | 30% | 10% | 5% | 2.55 |

Inspection of the percentage differences at scale points 2 and 3
(analogous to category "items") may seem large, but a difference
between means of .30 is not so impressive. The point to be made,
then, is that content categories, by showing results in percentages,
may overemphasize the subjective impression of sample differences.

## VI.  SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

In this study, rating scales did in general meet the expectations set out in the introduction. A considerable analytic advantage was gained in the ability to employ a multifactorial analysis of variance with the rating scales, so that the importance of location was shown relative to factors of sex, religion, and socioeconomic status. Results with the Conflict Scale indicated that there were location differences not due simply to the confounding of the latter factors. The Emphasis Scale analysis, however, demonstrated that there is one dimension, at least, where location is not the source of obtained differences. We are thus led to the conclusion that it is necessary to account for at least the cited demographic variables in analyzing cross-national data as used in this study. Rating scales and multifactorial analysis of variance present such a method.

To draw methodological conclusions, it is necessary to examine the problem from a new perspective.[13] The two methods compared in this study can best be thought of as two ways of ordering data; nominal ordering is represented by the "content category method", while the "rating scale method" represents an attempt to order categories on an interval scale. Now, we have found in the study that the "rating scale method" was not a complete substitute for the "content category method". Rather than consider this a deficiency of the former, we ought rather to consider the methodological explanations for this. First, the difficulties were partially due to the fact that the stories simply did not provide sufficient information relevant to some of the rating scale concepts. Second, the concepts themselves may have been inadequately formulated, e.g., definitions not precise enough, unidimensionality not established, etc.. And third, difficulties may have stemmed from

---

attempting to superimpose interval ordering on data not capable of this. These considerations indicate that subsequent research involving analysis of qualitative data ought to attend to the following:

1. <u>Obtaining relevant data.</u> Of necessity, this study was post hoc, and an attempt was made to superimpose scales for which the data were not originally intended. Ordinarily, however, the psychological variables of interest ought to be pre-established, then methods worked out until the data-gathering technique provides the necessary information. With incomplete stories, different stories could be developed to assess different psychological aspects.

2. <u>Determining scalar properties of the data.</u> Rather than assuming that the data are capable of interval ordering, it is desirable that this be established. Guttmann's method provides one way of determining whether categories form a scale, i.e. showing that the categories can be ordered sequentially. If not, then the categories may only be used in a naming procedure, which is essentially what we have done in the "content category method".

The kind of ordering made upon a set of categories seems directly to relate to the ease of development and the number of content aspects to which applicable. Thus, categories with mutually exclusive and mutually exhaustive items in the nominal-ordering "content category method" are applicable to fewer story aspects than categories without the restriction on the items. Rating scales, which make the further restriction that the items be sequentially ordered, are even more limited in application. Therefore, while this thesis posed the problem of choosing between two methods of content analysis, we must conclude that the "choice" is really established by the nature of the categories set forth. There is opportunity for choice only insofar as nominal ordering may be given to categories capable of interval ordering, but the reverse is not possible.

A final remark is directed to the statistical analysis of the ordered data. This study has presented a multifactorial method

applicable to interval-scale data, but an equivalent method for nominally ordered data does not seem yet available. At this time, then, we recommend use of interval scaling where possible, if multifactorial analysis is desired. This may help in interpreting nominally-ordered data, but the preferable solution to be sought is the development of a non-parametric multifactorial technique, directly applicable to this kind of ordering.

REFERENCES

Abrahamsen. D. Men, mind, and power. New York: Columbia Univer. Press, 1945.

Anderson, H. H. Children's values in Western Europe and the Americas. Paper read at Amer. Psychol. Ass., Chicago, Sept. 1960.

Anderson, H. H. & Anderson, Gladys L. Cultural reactions to conflict: a study by adolescent children in seven countries. In G.M. Gilbert (Ed.), Psychological approaches to intergroup and international understanding. Austin, Texas: Univer. of Texas, 1956.

Anderson, H.H., Anderson, G.L., Cohen, I.H., & Nutt, F.D. Image of the teacher by adolescent children in four countries: Germany, England, Mexico, United States. J. soc. Psychol., 1959, 50, 47-55.

Anderson, H.H., Anderson, Gladys, L., Daugherty, Mary Ann, Meyers, Maryanne L., Smith, L.J., & Mason, J.E. Authoritarian and democratic teacher-child relations in five countries as reported by adolescent children. Paper read at Mich. Acad. Sci., Arts, and Letters, Detroit, March, 1957.

Andregg, N.B. A critical study of graphic rating scales. Unpublished doctoral dissertation, Mich. State Coll., 1951.

Allport, G.W. Outlook of youth in ten countries. Acta Psychologica, 1955, 11, 211-212.

Auld, F.G. Influence of social class on personality test responses. Psychol. Bull., 1952, 49, 318-332.

Azimi, C. Sexual differentiation and identification in Iranian children of varying cultural backgrounds. Unpublished M.A. thesis, Mich. State Univer., 1960.

Bendig, A.W. The reliability of self-ratings as a function of the amount of verbal anchoring and of the number of categories on the scale. J. appl. Psychol., 1953, 37, 38-41.

Bendig, A.W. Reliability and the number of rating scale categories. J. appl. Psychol., 1954, 38, 38-41.

Bendig, A.W., & Sprague, Janine. Rater experience and the reliability of case history ratings of adjustment. J. consult. Psychol., 1954, 18, 207-211.

Bennett, C.A., & Franklin, N.L. Statistical analysis in chemistry and the chemical industry. New York: Wiley, 1954.

56

Brandt, A.E.  The analysis of variance in a 2 x 5 table with dis-
proportionate frequencies.  J. Amer. statist. Ass., 1933, 28,
164-173.

Brunswik, E.  Perception and the representative design of psy-
chological experiments.  Berkeley and Los Angeles:  Univer.
Calif. Press, 1956.

Buchanan, W., & Cantril, H.  How nations see each other.  Urbana,
Ill.:  Univer. Ill. Press, 1953.

Cattell, R.B.  The cultural functions of social stratification: II.
Regarding individual and group dynamics.  J. Soc. Psychol., 1945,
21, 25-56.

Cureton, E.E.  Note on the scaling of rankings when the number per
subject are unequal.  Psychometrika, 1952, 17, 397-399.

Davidson, Helen H.  Personality and economic background:  a study
of highly intelligent children.  New York:  King's Crown Press,
1943.

Dickens, Sara Lee, & Hobart, C.  Parental dominance and offspring
ethnocentrism.  J. soc. Psychol., 1959, 49, 297-303.

Duijker, H.C.J.  Comparative research in social science with special
reference to attitude research.  Int. soc. Sci. Bull., 1955, 7,
555-566.

Farlier, M.L.  The study of national character:  1955.  J. soc.
Issues, 1955, 11, 52-56.

Festinger, L., & Katz, D. (Eds.)  Research methods in the behavioral
sciences.  New York:  Dryden, 1953.

Frenkel-Brunswik, Else.  Differential patterns of social outlook
and personality in family and children.  In Margaret Mead &
Martha Wolfenstein (Eds.)  Childhood in contemporary cultures.
Chicago:  Univer. Chicago Press, 1955.

Frenkel-Brunswik, Else.  Parents and childhood as seen through
the interviews.  In T.W. Adorno, Else Frenkel-Brunswik, D.J.
Levinson, & R.N. Sanford, The authoritarian personality.  New
York:  Harper, 1950.

Fromm, E.  Escape from freedom.  New York:  Farrar & Rienhart, 1941.

Gillespie, J.M. & Allport, G.W.  Youth's outlook on the future.
Garden City, New York:  Doubleday, 1955.

Guilford, J.P.  Psychometric methods.  (2nd ed.)  New York:
McGraw-Hill, 1954.

Heber, R.F. A cross-cultural comparison of children's judgment of parent-child conflict in Germany, England, Finland, United States and Mexico. Unpublished M.A. thesis, Mich. State Coll., 1955.

Helson, H., Dworkin, R.S., & Michels, W.C. Quantitative denotations of common terms as a function of background. Amer. J. Psychol., 1956, 69, 194-208.

Himmelweit, Hilde T. Socio-economic background and personality. Int. soc. Sci. Bull., 1955, 29-34.

Jones, M.B. Religious values and authoritarian tendency. J. soc. Psychol., 1958, 48, 83-89.

Kempthorne, O. The design and analysis of experiments. New York: Wiley, 1952.

Klineberg, O. Tensions affecting international understanding: a survey of research. New York: Soc. Sci. Res. Council, 1950.

Levitov, Edith S. "The broken bicycle": a comparative study of story completions by adolescents in United States and German cities. Unpublished M.A. thesis, Geo. Washington Univer., 1960.

Lowie, R.H. Toward understanding Germany. Chicago: Univer. Chicago Press, 1954.

McClelland, D.C., Sturr, J.F., Knapp, R.H., & Wendt, H.W. Obligations to self and society in the United States and Germany. J. abnorm. soc. Psychol., 1958, 56, 245-255.

McGranahan, D.V. A comparison of social attitudes among American and German youth. J. abnorm. soc. Psychol., 1956, 41, 245-257.

Metraux, Rhoda. Parents and children: an analysis of contemporary child-care and youth guidance literature. In Margaret Mead & Martha Wolfenstein (Eds.), Childhood in contemporary cultures. Chicago: Univer. Chicago Press, 1955.

Miller, D.R. & Swanson, G.E. The changing American parent: a study in the Detroit area. New York: Wiley, 1958.

Mosier, C.I. A psychometric study of meaning. J. soc. Psychol., 1941, 13, 123-140.

Patterson, D.G. Rating. In D.H. Fryer & E.R. Henry (Eds.), Handbook of applied psychology. New York: Rinehart, 1950.

Robinson, Barbara M. A further cross-cultural comparison of children's judgment of parent-child conflict in Germany, Finland, England and the United States. Unpublished M.A. thesis, Mich. State Univer., 1955.

Rodnik, D. Postwar Germans. New Haven: Yale Univer. Press, 1948.

Rodnik, D. The Norwegians: a study in national culture. Washington: Public Affairs Press, 1955.

Rokeach, M. Political and religious dogmatism: an alternative to the authoritarian personality. Psychol. Monogr., 1956, 70 (425).

Schaffner, B. Fatherland: a study of authoritarianism in the German family. New York: Columbia Univer. Press, 1948.

Scheffe, H. The analysis of variance. New York: Wiley, 1959.

Selltiz, Claire, Jahoda, Marie, Deutsch, M., & Cook, S.M. Research methods in social relations. New York: Holt, 1959.

Sheppard, D. The adequacy of everyday quantitative expressions as measures of qualities. Brit. J. Psychol., 1954, 45, 40-50.

Siegel, S. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956.

Snedecor, G.W. Statistical methods applied to experiments in agriculture and biology. (5th ed.) Ames, Iowa: Collegiate Press, 1956.

Stevens, W.L. Statistical analysis of a non-orthogonal tri-factorial experiment. Biometrika, 1948, 35, 346-367.

Terhune, K.W. Story A-3 -- The Lost Money: special analysis of acting variables. Unpublished memorandum, 1959. (a)

Terhune, K.W. Story A-3 -- The Lost Money: special analysis of acting variables. Supplement. Unpublished memorandum, 1959.(b)

Terhune, K.W. Story A-3: Detailed analysis of acting variables - Mexico City. Unpublished memorandum, 1960.

The RAND Corporation. A million random digits with 100,000 normal deviates. Glencoe, Illinois: Free Press, 1955.

Walker, Helen M., & Lev, J. Statistical inference New York: Holt, 1953.

Wilson, K.V. A distribution-free test of analysis of variance hypothesis. Psychol. Bull., 1956, 53, 96-101.

Yates, F. The principle of orthogonality and confounding in replicated experiments. J. agricult. Sci., 1933, 23, 108-145.

Yates, F. The analysis of multiple classifications with unequal numbers in the different classes. J. Amer. statist. Ass., 1934, 29, 51-66.

APPENDICES

Anderson Imcomplete Stories:  Series B

| | Age at last | | |
Birthdate:_____ _____Birthday:_____Boy:_____Girl:____

School:_____Grade:_____Church:_____

Occupation of father:_____Occupation of mother:_____

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

1. Mary's grandmother gave her a beautiful vase for her thirteenth birthday.  One day Marys mother comes in with flowers which she puts in the vase.  She places the vase in the window, though she knows the window ledge is too narrow for this vase and that it might fall off.  As her mother is dusting, she bumps the vase and it crashes to the floor and breaks.  While Mary's mother is in the kitchen getting a cloth, Mary returns from school and finds the vase in pieces on the floor.

> What does Mary do?          What does the mother say?
> How do they both feel about it?

Think about these questions and __finish__ this story quickly with a few sentences.

2. John worked hard and bought a new bicycle with his money. One Saturday, while John is playing with friends, John's father decides to ride this bicycle to the store to buy cigarettes. He leaves the bicycle in the street. When he comes out of the store he finds the front wheel bent and some of the paint on the frame badly scrapped. No one is around. The father could still ride the bicycle home.

What does the father do?                    What does John do?
How do they both feel about it?

Think about these questions and <u>finish</u> the story quickly with a few sentences.

3. The teacher has promised the class that if they work hard during the last month of the term they will have one day off to make a special trip. Several times during the last month the teacher talks about this trip. The children want to make this trip and they work very hard. Now it is the last week of the term and there is no time for a trip.

What does the teacher do?
What do the children think, and how do they feel about it.

Think about these questions and <u>finish</u> this story quickly with a few sentences.

4. Herbert received for his thirteenth birthday a handsome camping axe. It is sharp and has a strong leather case. While Herbert is at school his four-year-old brother, Billy, sees the axe, looks at it a long time, picks it up, puts it back, and finally takes it outdoors with him to play. Billy does not take the case off. He sings to himself as he walks about the garden, tapping the axe gently against a tree, a post, and the pavement. Herbert comes home from school, finds the axe in its leather case with some other toys. But he sees that the axe has cut through the leather case and the blade is chipped and blunted.

What does Herbert do?
How does Herbert feel about it?

Think about these questions and finish the story quickly with a few sentences.

5. Kate, aged 13 years, has a little four-year-old sister, Clara. When Kate comes home from school Clara often wants to play with her and follows her and her older girl friends around. One day Clara took a very pretty new dress from Kate's wardrobe, put it on herself, and looked in the mirror. It was too long and hung to the floor. She gathered up the skirt in her arms and went out of the house for a walk. Kate came home from school, found her new dress crumpled on a chair. The skirt had been stepped on and dragged in the dirt.
     Clara said, "I wore your dress."

What does Kate do?
What do Kate and Clara think about it and how do they feel?

Finish the story in a few sentences.

Michigan State University                           Psychology-Creativity Research

ANDERSON INCOMPLETE STORIES - MASTER LOCATION LIST*

| Location | Ab-rev. | IBM Code | Date of Admin. | Number of Children Used in Analyses | | | | |
|----------|---------|----------|----------------|-----------|-----------|--------|-------|---------|
| | | | | No of 7th Grade | No of 4th Grade | Other Grades | Total No. | Series Avail-able |
| Karlsruhe Germany | Kar | 01 | Sep '52 | 1208 | --- | --- | 1208 | A |
| Braunschweig Germany | Brn | 02 | Dec '53 | 187 | 383 | 147(10th) | 717 | A&B |
| Hamburg Germany | Ham | 03 | Jan '54 | 524 | 575 | --- | 1099 | A&B |
| Munich Germany | Mun | 04 | Feb '54 | 350 | 498 | --- | 848 | A&B |
| Flensburg-Murwik Germany | F-M | 05 | Jan '54 | 57 | --- | --- | 57 | A |
| Broitzem Germany | Brz | 06 | Dec '53 | 28 | 30 | --- | 58 | A |
| Birmingham England | Bhm | 07 | Mar '54 | 402 | 285 | --- | 687 | A&B |
| Drammen Norway | Drm | 08 | Mar '54 | 225 | 256 | --- | 481 | A&B |
| Stockholm Sweden | Stk | 09 | Mar '54 | 186 | 235 | --- | 421 | A&B |
| Helsinki Finland | Hel | 10 | Apr '54 | 220 | 194 | --- | 414 | A&B |
| Mexico City Mexico | Mex | 11 | May '53 | 1254 | --- | --- | 1254 | A |
| Tecomitl Mexico | Tcl | 12 | May '53 | 110 | --- | --- | 110 | A |
| Tepoztlan Mexico | Tpz | 13 | May '53 | 99 | --- | --- | 99 | A |
| Tzintzuntzan Mexico | Tzn | 14 | May '53 | 17 | --- | --- | 17 | A |
| Knoxville Tennessee | Knx | 15 | Apr '53 | 226 | --- | --- | 226 | A |
| Benton Harbor Michigan | BnH | 16 | Jan '57 | 299 | --- | --- | 299 | A&B |
| San Juan Puerto Rico | SuJ | 17 | Dec '56 | 535 | --- | --- | 535 | A&B |
| Girl's Training School Adrian, Michigan | Adr | 18 | Jul '56 | --- | --- | 166(older mixed) | 166 | A |
| Chiconcuac Mexico | Ccc | 19 | Oct '59 | --- | --- | 63(mixed) | 63 | A |
| Rio de Janeiro Brazil | Rio | 20 | Aug '59 | | | | | A&B |
| Iowa City Iowa, Exper. | Ice | 21 | Jul '60 | --- | --- | 80(6th) | 80 | A-5 |
| Iowa City Iowa, Control | Icc | 22 | Jul '60 | --- | --- | 89(6th) | 89 | A-5 |
| Munich Germany | MuS | 23 | Mar '60 | | | | | A&B |

*Revised on 1 August, '60 from Master Location List of 8 May, '58

DJE
KWT

## ANDERSON INCOMPLETE STORIES

Harold H. Anderson and Gladys L. Anderson
Department of Psychology, Michigan State University
East Lansing, Michigan

### Directions for Administering

Here are six short stories. Boys and girls in other American schools as well as in Germany, Mexico and England have also written these stories. Read what happens in the story, then write what, in your opinion, happens next. Use your imagination. Write clearly, but do not try to write beautifully. Try to write as quickly as you can. If you make a mistake, strike out the word and continue immediately.

You will not be graded for these stories. Your teacher, the principal, and other teachers in this school will not even get to see these stories. We do not want you to put your name on the paper. Write honestly what you really think. If your name is not on the paper, you may even write something which you would not tell to a teacher or to anyone else.

I will now hand out the forms. You need not look at each other's papers as there are no right or wrong answers. You each have enough imagination to write for yourselves. (Hand out the papers. The teacher does not participate.)

On the top of the first page please write the date of your birth. Write the month, the day and the year, like this (write on the board) June 4, 1943. Then write how old you were on your last birthday (write 13). Now make a cross after boy or girl. Write the name of your school and give your school grade. Write what church you attend, or if you do not attend church, write none. When you write the occupation of your father, do not write the name of the company, but the kind of work he does, like carpenter, laborer, bookkeeper. If your mother works away from home, write what she does. Otherwise write home.

If you read the first story you will see that there are some questions at the end. Please think about these questions, but do not try to answer them one by one. These questions are put there to help you finish the story. When you finish the first story, continue with the next story without delay. If I (test administrator) go through the aisles I am not trying to look at your paper and read what you have written. I will only see how far you have proceeded.

Try to finish all six (five) stories. You will have the entire period. Do you have any questions? If you cannot read a word, or if you have a question, or need a pencil, please raise your hand and I will then come to your seat. And now are you ready? If so, ready, start.

January 1957.

### Levels of Socio-Economic Status
### Classified by Father's Occupation*

S.E.S.-1:   Upper middle (professional)
>Professional, technical, and kindred workers
>Managers, officials, and proprietors except farm

S.E.S.-2:   Lower middle ("white collar")
>Clerical and kindred workers
>Sales worker
>Public school teachers
>Nurses
>Draftsmen
>Lab technician
>Managers of small offices
>City firemen
>Policemen
>Investigators
>Sheriffs

S.E.S.-3:   Upper lower and lower lower ("blue collar")
>Craftsmen, foremen, and kindred workers
>Service workers
>Laborers
>Farmer, farm hand

### Arbitrary placement of vaguely stated occupations

>S.E.S.-2:   "Employee", "Merchant", "Musician", "Super-
>visor", "Salesman"
>S.E.S.-3:   "Worker", "Foreman", "Skilled worker"

---

* The classification scheme is that devised by Miller and Swanson, based on the 1950 United States Census Bureau placements and modified slightly according to a study of prestige by the National Opinion Research Center. See: Miller, D.R., & Swanson, G.E. *The Changing American Parent*. New York: Wiley, 1958.

## Outline of Coding Scheme for Story B-4
(Categories begin at 23 for I.B.M. purposes)

21 <u>Underlying cause of situation</u> (CHECK)
    1 No indication of fault
    2 Billy's fault
    3 Herbert's fault
    4 Herbert's and Billy's fault
    5 Mother's fault, immediate
    6 Mother's fault, past

22 <u>Herbert's initial action</u> (CHECK)
Information seeking; no assumption as
to guilt. (Suspicion accepted.)
    1 Asks Billy open question
    2 Asks mother or father open question
    3 Asks general question of anyone, or vague
Not information seeking; assumes B. respons.
    4 Approaches Billy
    5 Approaches mother or father
    6 Other, including assumptive vague
Neither info. seeking, nor assume guilt
    7 No specification or indic. of assumption

23 <u>Communication by Billy re. axe</u> (CHECK)
    1 No comm. or vague as to comm. re. axe
    2 Billy tells truth re. axe
    3 Billy lies, deceives re. axe

24 <u>Herbert's primary actions re. Billy</u> (CHECK)
    1 No primary action, vague or ambiguous
    2 Advisory, explanatory toward Billy
    3 Pardons, forgives, excuses Billy (overtly)
    4 Punitive toward Billy (scolds, threatens, etc.)
    5 Tells on Billy

25 <u>Herbert's primary actions re. axe</u> (CHECK)
    1 Omits restoration, vague or ambiguous
    2 Effects restoration independently
    3 Effects restoration dependently; asks someone
    4 Demands, orders Billy to restore
    5 Mutual arrangement with Billy
    6 Restoration without Herbert

26 Herbert's secondary actions (CODE)
   1 Forgives, consoles, apologizes to Billy
   2 Prohibits Billy from touching axe
   3 Makes axe inaccessible to Billy
   4 Realizes should hide axe; no action indic.
   5 Gives axe to Billy
   6 None of the above, or vague

27 Herbert's unfulfilled intentions, desires
re. punishment (CHECK)
   1 None specified
   2 Punish by self; decides against, just doesn't
   3 Punish by self; prevented, interrupted
   4 Have m. or f. punish; decides against
   5 Have m. or f. punish; parents don't
   6 Punishment, gen'l; unclear why not

28 Punishment of Billy: source (CHECK)
   1 No punishment administered
   2 Unclear as to fact of punishment
   3 Punishment by Herbert
   4 Punishment by mother
   5 Punishment by father
   6 Punishment by mother and father
   7 Punishment by Herbert and parent(s)
   8 Punishment by unclear, unspec. source

29 Punishment of Billy: mode (CHECK)
   1 No punishment mentioned
   2 Verbal, non-constructive
   3 Verbal, constructive
   4 Physical
   5 Material deprivation, revenge
   6 Material deprivation, repayment
   7 Activity deprivation
   8 Verbal and physical
   9 Verbal and deprivation
  10 Physical and deprivation
  11 Verbal, physical, and deprivation
  12 Vague as to mode

30 Billy's responses to Herbert (CHECK)
   1 Agrees with Herbert, reforms, etc.
   2 Apologizes, asks forgiveness
   3 Voluntary restoration of axe
   4 Consoles, sympathizes with Herbert
   5 Argues, resists Herbert
   6 Runs to mother or father
   7 Just runs away (into house, etc.)
   8 None of the above, or vague

31 <u>Parental inclusion</u> (CHECK)
    1 Not included (not mentioned)
    2 Herbert brings in
    3 Billy brings in
    4 Other, or vague

32 <u>Actions by parents</u> (CODE)
    1 Parents not included
    2 Take no action; none specified
    3 Seek or communicate info; answer ques.
    4 Advise, suggest, explain, mediate
    5 Effect restoration of axe and/or case
    6 Console Billy or Herbert
    7 Overtly reject Billy's behavior
    8 Overtly reject Herbert's behavior
    9 Other, including vague

33 <u>Feelings: Herbert re. damage to axe</u> (CHECK)
    1 None specified
    2 Any of: annoyed, cross, irritated, etc.
    3 Any of: angry, mad, furious, etc.
    4 Any of: sad, unhappy, upset, cries, etc.
    5 Angry/annoyed AND sad
    6 Other

34 <u>Feelings Herbert re. other than damage to axe</u> (CHECK)
    1 None specified
    2 Any of: guilt, shame, regrets actions, etc.
    3 Any of: sadness, sorrow, cries, etc.
    4 Any of: happy, glad, satisfied, etc.
    5 Other

35 <u>Feelings: Billy</u> (CHECK)
    1 None specified
    2 Guilt, shame, blushes, etc.
    3 Unhappy, forlorn, sad, cries, etc.
    4 Fear, scared, trembles, etc.
    5 Sorry, regretful, repentant, etc.
    6 Happy, proud, etc.
    7 Angry, mad, etc.
    8 Other

36 <u>Clues: Herbert-Billy relationship</u> (CODE)
    1 Herbert: concern for Billy's safety
    2 Herbert: understanding re. Billy's intent
    3 Billy: sympathy toward Herbert
    4 Either or both boys: affection, comradeship
      expressed
    5 Either or both boys: resentment, jealousy
    6 Billy omitted from the story
    7 None of the above

37 Restoration of axe  (CHECK)
    1 No action re. restoration, or vague
    2 New axe thema
    3 Repair axe thema


## HOW TO CODE

CHECK category - Always check one, but only one
  item.  (Items are mutually exclusive and mutually
  exhaustive.)

CODE category - Always check at least one item.
  (Items are mutually exhaustive, not mutually ex-
  clusive.)

You will be given a number of stories which are children's completions to story B-4 of the Anderson Incomplete Stories. Your task as a judge will be to rate the story completions on each of four rating scales, which will be explained shortly.

The stories will be given to you in random order and you are to make your ratings in that order. You are being provided with scoring forms on which to record your ratings. You will find that each story completion is numbered, and you are also to record the story number on the scoring form in the place provided.

Each scale will be divided into eleven divisions, a number found suitable in a pilot study on this material. When you rate a story on a scale, you are to mark with an "X" that division which you think is the most appropriate position on the continuum. Occasionally a judge feels that a rating should be half-way between two divisions. In this case, the judge may mark his "X" on the dividing line between two divisions.

To guide you in the meaning of the various scale positions, you will find attached to these instructions detailed descriptions of the scales, with example stories at different points along each continuum. These stories have been agreed upon by judges in a pilot study, but you need not agree absolutely with their ratings. The example stories are provided merely as auxiliary guides for you.

On three of the four rating scales you will find a box where you may mark "cannot rate". The box may be checked by the rater for those story completions where the rater feels that the scale cannot appropriately be applied, or he feels too uncertain to give a rating. You are urged to be conscientious and rate when you can, but do not rate when it isn't warranted. (When in doubt as whether to give a rating, be liberal and assign a rating.) There will be one scale which does not have a "cannot rate" box, and every story completion must be rated on this scale.

To facilitate your understanding of the story completions, the original Incomplete Story B-4 which the children completed is presented below.

> Herbert received for his thirteenth birthday
> a handsome camping axe. It is sharp and has a strong
> leather case. While Herbert is at school his four-
> year-old brother, Billy, sees the axe, looks at it
> a long time, picks it up, puts it back, and finally
> takes it outdoors with him to play. Billy does not
> take the case off. He sings to himself as he walks

about the garden, tapping the axe gently against
a tree, a post, and the pavement. Herbert comes
home from school, finds the axe in its leather
case with some other toys. But he sees that the
axe has cut through the leather case and the blade
is chipped and blunted.

The scales which you will use will now be described fur-
ther. As the scoring forms will contain only abbreviated descrip-
tions, you may find it helpful to refer back to these extended des-
criptions from time to time.

## Soale 1: Dominance Scale

This scale pertains to Herbert's behavior toward Billy. A
rating should be given only where there is direct, face-to-face in-
teraction between the two boys; otherwise check "cannot rate".
(Thus, Herbert simply reporting to a parent about Billy is not con-
sidered in rating on this scale.)

By "dominance" we mean here a state of relations or inter-
actions between two individuals in which one individual controls the
other's behavior to a certain degree, this control being indepen-
dent of and without regard to the desires of the person being con-
trolled. It is an enforcement of behavior.

## Scale 2: Conflict Scale (Direction of Herbert's behavior)

Here the rater evaluates Herbert's overall behavior in the
story and where it seems to be leading in terms of interpersonal re-
lations with Billy. Both overt and covert behavior are to be con-
sidered. The judge should keep in mind the effects of behavior
earlier in the story as well as the last behavior mentioned. Every
story is to be rated on this scale, except those in which Herbert
is omitted from the story (rare). Where the coder is unsure of the
direction, he should simply score "neutral" on the scale.

## Scale 3: Emphasis Scale

On this scale the judge records his assessment of the re-
lative weight or emphasis given to material things (the axe, its
damage, repair, etc.) as opposed to people (concern for Billy,
feelings toward Billy, punishment of Billy, teaching of Billy, etc.)
The rater should consider the motivations involved, the goals being
sought. Thus in evaluating Herbert's behavior, the rater should
ask himself, "What is Herbert trying to accomplish? Are his aims
to do something about the axe or to do something with people? Are
his feelings focused on the axe or on Billy?".

As an example, suppose that Herbert scolds Billy: he may tell Billy, "Don't you touch my things again!", with the emphasis on protection of material possessions, or he may say "You are a disagreeable brat", with the emphasis on condemning Billy as a person.

In coding a story on this scale, the rater should consider primarily content in assessing the main concern or emphasis. Thus, Herbert's feelings should weigh heavily in the rating. Emphasis is also to be assessed by the amount of discussion or narrative devoted to persons vs. things, but this should be given secondary importance.

Every story is to be rated on this dimension. In marking the scale, the coder is to divide the total length in proportion to the relative emphasis on persons vs. things. The left-hand side represents the amount of emphasis on persons, the right side the emphasis on things. Caution: this is easily confused in the rating process; remember -- left portion = persons, right portion = things. The following example illustrates this.

INCREASED EMPHASIS ON PERSONS

X

25% of emphasis is on persons

75% of emphasis is on things
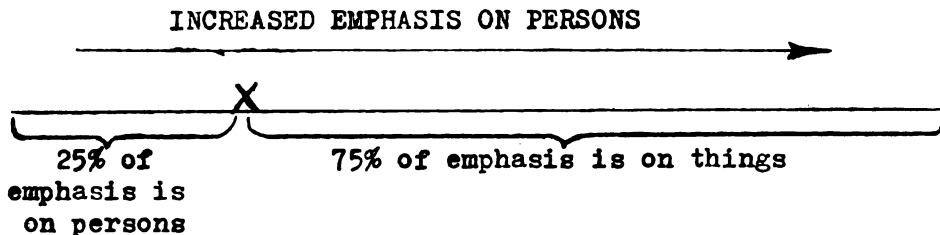
## Scale 4: Punishment Scale

Here the rater is to assess the degree of severity of punishment administered to, or intended for Billy, regardless of who administers or plans the punishment. (Caution: there are to be no inferences here. The story must actually state what the punishment is to be.) "Punishment" is to be interpreted broadly here, including castigation (tongue-lashing, scolding) chastisement, and disciplinary action. It thus includes both corporal and psychological punishment.

When the punishment is not explained, or is not clear, check "cannot rate".

## Supplement to Rating Scale Instructions

1. **Halo effect** -- although we expect the scales to be correlated, try to evaluate each scale for its own merits.

2. **Punishment scale** -- In general, code only that punishment which appears in the story or is definitely intended. Threats do not count as punishment. An exception to the rule here will occur when there is definite intent to punish Billy, but the intender (Herbert, parent) voluntarily changes his mind before administering the punishment. If the intender is only prevented from administering the punishment because of interference from an outside source, then rate as if the punishment actually occurred.

   While single punishments will generally be rated as indicated by the cues on the scale, multiple punishments should generally be considered as more severe than any one of the punishments administered.

   The cue on the scale marked "deprivation" refers to deprivation qua punishment, and is to be considered different from restitution (e.g., payment for a new axe). Examples of deprication would be destroying of Billy's toys, being sent to bed early (activity deprivation), etc.

3. **Emphasis scale** -- Imperative to keep in mind, to consider the aims, goals of any given activity. Although a certain interaction may be interpersonal, its intent or concern may center around the damaged axe, and therefore the emphasis would be on things. It seems that more deliberation is necessary before assigning a rating on this scale than for the others.

   Use the emphasis scale as a five-point scale.

4. **Inferences** should be reduced to a minimum. Go by the facts of the story. Rather than make a rating in which you have little confidence, score "cannot rate". This is somewhat in contradiction to the original instructions. Let us say this: if you are a little doubtful as to the exact position of a rating, go ahead and score it. If, however, you can see that the story is ambiguous and could be interpreted in two radically different ways, then better to leave well enough alone and score "cannot rate". And of course, always read the stories carefully so as not to miss important (and perhaps subtle) points.

5. **Cues** -- Become familiar with the more explicit cues on the example scales in the instructions. This is especially necessary on the direction of behavior scale -- know the distinction between the scale divisions.

6. <u>Resuming set</u> -- When resuming rating after having some time off,
probably is a good idea to review some of your previous ratings
to re-establish your set or adaptation level. You might even re-
score some of the previous stories to check your set.

## Dominanee Scale:  Cues and Examples

**Total dom. of Billy by Herb. Complete rule, control over B's behav. En- forced obed- ienee. B's de- sires not con- sidered.**

Herbert kills his brother.

He will whip Billy thoroughly and make him make up for it.

He hits little Billy and forbids him ever to touch his toys again.

Herbert says, "Billy, you have to buy me a new axe." Billy got his savings bank. The money was just enough. Billy was very embarrased about the matter and he handed the money to Herbert.

Herbert looked at the hatchet carefully and he wondered how he could get it into order again. He could sharpen the blade again but the cap could not be sewed. He hollered at his brother and he beat him properly. He wished for a new hatchet for Easter because he had had a good report card he received one too.

Herbert gives his brother a good one to remem- ber this time. He throws all kinds of things at his head. Little Billy thinks naturally what is the matter with him because he doesn't know at all that you have to take the cover off. He is very mueh insulted.

Herbert scolds his brother Billy but he did not get a new one again.

**INCREASING DOMINANCE**

**No dom. of B. by Herb. B's. evert behav. under no con- trol by Herb. B. free to fol- low own de- sires.**

He goes to Billy and ask him whether he has had the axe. Billy answered yes I have played with it. How do you come to play with my things? You have ruined it. I did not want to do that said Billy and started to cry. As he sees his brother cry like this certainly his heart softens and he forgives him also gives him the axe because he cannot use it any longer like this and they conclude peace.

## Conflict Scale: Cues and Examples

**Herb's behav. tending toward ult. strife, enmity.H.hates desires revenge.**

*INCREASING CONFLICT*

Herbert asked his little brother and finds out that he's the guilty one so he wants to get even. He destroys Billy's prized teddy bear.

**Unassuaged anger,bitterness by H.Desires intends "get even".**

He will whip Billy thoroughly and make him make up for it.

**Very angry, bitter scolding of B.**

Herbert wanted to revenge himself.Therefore he went to the brother and scolded him considerably. And he ran to his mother but there he also was scolded.And thus the matter with the axe was finished.

**H. "tattles".**

He tells his mother that Billy had cut the leather case and blunted the axe.Well Billy has a job he can buy you a new case and have the axe sharpened "okay mom"

**Neutral**

**Questions B. objectively,no malice.**

*INCREASING HARMONY*

Herbert ask his mother who had his axe.Mother said that she saw Billy look at it.He ask Billy did he take it.Billy said yes then his mother told Herbert that she would get him another one that is not to leave it lying around.

**Forgiveness, any anger is brief.No hostility.H. accepts situation**

Herbert don't be angry.I still have 20 pennies. You can buy yourself a new axe with it.Then Herbert had to laugh and he couldn't be angry at Billy anymore.He wished for a new axe for Christmas.

**H. understands B's age a factor,not angry.**

Herbert thinks to himself my little brother is still too little to understand this.He lets it lie there and does not look at it anymore.He thought later I have it fixed when I am bigger.

**H.shows understanding,seeks prevention,yet satisfies B.**

**Toward complete harmony,comradeship,good will.**

Herbert sees Billy and he asks him why he has taken the axe then Billy says,because he wants to have one like that too then Herbert said, why shouldn't he get one and he made one for his brother from wood because he liked to hit against stone and iron with it and because then he can't hurt any other child with it either.

## Emphasis Scale:   Cues and Examples

**Almost all emphasis on persons**

Herbert ask about it and his brother said he had done it.Herbert was sad and was going to slap his brother when his mother came in and ask what was going on.Herbert told her and little Billy started to cry and Herbert said don't cry Billy  I will let it go this time.Little Billy said it won't happen again.

**More on persons than things.**

"Actually, I am to be blamed myself,"thinks Herbert. "If I had put away the axe, Billy would not have gotten hold of it."

**Equal emphasis**

(1) Herbert told his parents and they found out how it happened and Herbert received a new axe. They also showed their displeasure toward Billy.

(2) Herbert goes and tells his father about it and says that probably his 4 year old brother was playing with (it) but didn't mean to harm the case or the axe.Herbert's father says he will buy him a new case and perhaps a new axe for his fourteenth birthday.

**More on things than persons**

Herbert first gives him a box then he tells his father.He thinks now I could play with it so nice but now it is broken.I hope I will get a new one.

**Almost all emphasis on things.**

Herbert went and ask his mother who had his axe. Mother said that she saw Billy look at it.He went and ask Billy did he take it.Billy said yes then his mother came in and told Herbert that she would get him another one that is not to leave it lying around.

## Punishment Scale: Cues and Examples

**Extreme pun-ish. Psychol. or physic-ally brutal.**

Herbert takes the axe and uses it on his brother and then dumps his brother in the river and the axe because he wants to get rid of the evidence.

**Physical beating.**

Herbert tells his parents and Billy gets a beating. Herbert is very mad about the broken axe and cover. Herbert has it sharpened again and has the cover sewed again.

**Deprivation, slapping, hitting.**

Herbert wrestles with his brother. When the mother came from work Herbert told her what happened with the axe. When the mother learned about it he (the little brother) was not allowed to take the trip with them.

**Spanking**

**Angry, scolding, castigation.**

Herbert is very angry. He doesn't want any explanations from any one. His father says he will sharpen the axe for him, but Herbert doesn't listen. He spanks his little brother, so therefore his father takes the axe away from him.

**Simple scolding.**

Herbert scolds Billy and then he puts it into a safer place where Billy can't get at it. He feels it is not nice that Billy has gone at it and has taken it for play.

**No punish. administ.**

Herbert sees Billy and he asks him why he has taken the axe then Billy says, because he wants to have one like that too then Herbert said, why shouldn't he get one and he made one for his brother from wood because he liked to hit against stone and iron with it and because then he can't hurt any other child with it either.

## Sample Story Rating Form

Story no. _____

Rater's Initials _____ Date _____

Can't Rate ☐

**Dominance by Herbert**

NO DOM. ———————————————— TOTAL DOMINANCE

Can't Rate ☐

**Direction Herb's behavior**

TOWARD COMPLETE HARMONY | UNDERSTANDING POSITIVE ACTIONS | FORGIVES ACCEPTS SITUATION | NEUTRAL | BITTER SCOLDING CASTIGATION | HITTING BEATING BITTERNESS | TOWARD ULTIMATE STRIFE CONFLICT

**Emphasis in story**

ALMOST ALL ON THINGS | MORE ON THINGS THAN PERSONS | EQUAL | MORE ON PERSONS THAN THINGS | ALMOST ALL ON PERSONS

Can't Rate ☐

**Severity of punishment**

NONE | SIMPLE SCOLDING | ANGRY SCOLDING CASTIGATION | SPANKING SLAPPING HITTING | DEPRIVATION | BEATING | EXTREME EXCESSIVE SEVERITY, BRUTAL

APPENDIX H


Results of the Content Category Analysis

In the following tables are presented the results for all
the categories of the coding scheme for story B-4.* Denominators
for the percentages are specified as either N or n. N designates
the size of the total location sub-samples, while n is used where
percentages were calculated over a reduced portion of the sub-
samples.


Table H-1. Underlying Cause of Situation (Category 21)

| | | Item(s) | | |
| | | 1<br>No one's<br>fault | 2<br>Billy's<br>fault<br>entirely | 3,4<br>Herbert<br>at<br>fault |
| | N | | | |
| Drammen | 101 | 16.8% | 70.3% | 12.9% |
| Munich | 49 | 22.4% | 67.3% | 10.2% |
| Hamburg | 50 | 14.0% | 76.0% | 10.0% |
| Benton Harbor | 99 | 21.0% | 61.0% | 18.0% |

$\chi^2 = 5.3$   Not Significant


Table H-2. Herbert's Initial Action (Category 22)

| | | Item | | |
| | | 1,2,3<br>No assump.<br>info.<br>seeking | 5,6<br>Assumes<br>Billy<br>guilty | 7<br>No<br>info. |
| | N | | | |
| Drammen | 101 | 26.7% | 63.3% | 9.9% |
| Munich | 49 | 30.6% | 61.2% | 8.2% |
| Hamburg | 50 | 24.0% | 70.0% | 6.0% |
| Benton Harbor | 99 | 31.3% | 59.6% | 9.1% |

$\chi^2 = 2.0$   Not Significant


* Results of categories 34 and 35 were omitted because of very low
frequencies.

Table H-3.  Communication by Billy Regarding Axe (Category 23)

| | n<br>($\Sigma$Items 2&3) | Item | |
| | | 2<br>Tells<br>truth | 3<br>Tells<br>lie |
| --- | --- | --- | --- |
| Drammen | 23 | 82.7% | 17.3% |
| Munich | 12 | 58.4% | 41.6% |
| Hamburg | 18 | 72.2% | 27.8% |
| Benton Harbor | 17 | 88.3% | 11.7% |

$X^2$= 4.2   Not Significant

Note:  Item 1 of category 23 is "No communication by Billy"

Table H-4.  Herbert's Primary Actions Regarding Billy
(Category 24)

| | n<br>($\Sigma$ items 2-5) | Item(s) | | |
| | | 2,3<br>Advises,<br>Pardons | 4<br>Punishes<br>Billy | 5<br>Tells on<br>Billy |
| --- | --- | --- | --- | --- |
| Drammen | 52 | 23.1% | 71.2% | 5.8% |
| Munich | 29 | 13.8% | 82.7% | 3.5% |
| Hamburg | 32 | 6.3% | 68.7% | 25.0% |
| Benton Harbor | 59 | 13.6% | 45.8% | 40.6% |

$X^2$= 30.4    P <.001

Note:  Item 1 of category 24 is "No action by Herbert"

Table H-5.  Herbert's Primary Actions Regarding Axe
(Category 25)

| | N | Item(s) | |
| | | 1,6<br>Not re-<br>stored by<br>Herbert | 2,3,4,5<br>Axe re-<br>stored by<br>Herbert |
| --- | --- | --- | --- |
| Drammen | 101 | 77.2% | 22.8% |
| Munich | 49 | 91.8% | 8.2% |
| Hamburg | 50 | 76.0% | 24.0% |
| Benton Harbor | 99 | 91.9% | 8.1% |

$X^2$= 12.9    P <.01

Table H-6. Herbert's Secondary Actions (Category 26)

| | N | Item(s) | |
| --- | --- | --- | --- |
| | | 3,4 Realizes Billy's youth | 1,2,5,6 Other |
| Drammen | 101 | 16.8% | 83.2% |
| Munich | 49 | 14.3% | 85.7% |
| Hamburg | 50 | 14.0% | 86.0% |
| Benton Harbor | 99 | 13.1% | 86.9% |

$X^2 = 0.6$    Not Significant

Table H-7. Herbert's Unfulfilled Intentions of Punishment (Category 27)

| | N | Item(s) | |
| --- | --- | --- | --- |
| | | 1 No unful. Intentions | 2-6 Unfulfilled Intentions |
| Drammen | 101 | 93.1% | 6.9% |
| Munich | 49 | 98.0% | 2.0% |
| Hamburg | 50 | 84.0% | 16.0% |
| Benton Harbor | 99 | 94.9% | 5.1% |

$X^2 = 9.2$    $P < .05$

Table H-8. Punishment of Billy: Source (Category 28)
a. Presence or absence of punishment

| | N | Item(s) | |
| --- | --- | --- | --- |
| | | 1 No Punishment | 2-8 Punishment |
| Drammen | 101 | 53.5% | 46.5% |
| Munich | 49 | 44.9% | 55.1% |
| Hamburg | 50 | 42.0% | 58.0% |
| Benton Harbor | 99 | 59.6% | 40.4% |

$X^2 = 5.4$    Not Significant

b. Source of punishment, when specified

| | n (Σ items 3-7) | Item(s) | |
| | | 3 By Herbert | 4,5,6,7 By parents |
| --- | --- | --- | --- |
| Drammen | 26 | 87.2% | 12.8% |
| Munich | 26 | 96.2% | 3.8% |
| Hamburg | 39 | 69.3% | 30.7% |
| Benton Harbor | 38 | 68.4% | 31.6% |

$$X^2 = 10.5 \quad P < .05$$

Note: Item 1 is "No punishment", items 2 & 8 are "not clear"

Table H-9. Punishment of Billy: Mode (Category 29)

| | n (Σ items 2-11) | Item(s) | | | |
| | | 2,3 Verbal | 4 Physical | 5,6,7 Deprivation | 8,9,10,11 Multiple |
| --- | --- | --- | --- | --- | --- |
| Drammen | 45 | 24.4% | 51.1% | 15.6% | 8.9% |
| Munich | 26 | 30.7% | 57.7% | 3.8% | 7.7% |
| Hamburg | 36 | 30.5% | 46.2% | 16.7% | 5.6% |
| Benton Harbor | 34 | 23.5% | 47.1% | 20.6% | 8.8% |

$$X^2 = 4.3 \quad \text{Not Significant}$$

Note: Item 1 is "No punishment", item 2 is "vague"

Table H-10. Billy's Responses to Herbert (Category 30)

| | n (Σ Items 1-7) | Items | |
| | | 1-4 Cooperative | 5-7 Non-cooperative |
| --- | --- | --- | --- |
| Drammen | 14 | 78.7% | 21.3% |
| Munich | 5 | 40.0% | 60.0% |
| Hamburg | 6 | 100.0% | 0% |
| Benton Harbor | 6 | 67.7% | 33.3% |

$$X^2 = 5.5 \quad \text{Not Significant}$$

Note: Item 8 is "None"

Table H-11.  Parental Inclusion (Category 31)

a.  Fact of inclusion

| | N | Item(s) 1 Parents not incl. | 2,3,4 Parents included |
|---|---|---|---|
| Drammen | 101 | 62.4% | 37.6% |
| Munich | 49 | 61.2% | 38.8% |
| Hamburg | 50 | 50.0% | 50.0% |
| Benton Harbor | 99 | 39.4% | 60.6% |

$$X^2 = 12.4 \qquad P < .01$$

b.  How parents brought into story

| | n ($\Sigma$ Items 2,3,4) | Item(s) 2,3 Children bring in | 4 Other |
|---|---|---|---|
| Drammen | 38 | 65.8% | 34.2% |
| Munich | 19 | 73.7% | 26.3% |
| Hamburg | 25 | 64.0% | 36.0% |
| Benton Harbor | 60 | 73.3% | 26.7% |

$$X^2 = 1.2 \qquad \text{Not Significant}$$

Table H-12.  Actions by Parents (Category 32)

In this category, items are not mutually exclusive, so chi-squares must be computed on individual items.  Only in items 4, 5, 7, and 8 was this considered necessary.  This information is presented in a single table below.  All chi-squares were computed by comparing frequencies of the item in question with the remaining frequency.

| | Item 2 No action | 3 Provide info. | 4 Advise | 5 Restore Axe | 6 Console | 7 Reject Billy | 8 Reject Herbert | 9 Other |
|---|---|---|---|---|---|---|---|---|
| Drm | 15.8% | 15.8% | 23.7% | 31.6% | 2.7% | 18.4% | 15.8% | 5.2% |
| Mun | 36.8% | 21.0% | 5.3% | 31.6% | 0% | 10.5% | 5.3% | 0% |
| Ham | 0% | 20.8% | 20.8% | 45.8% | 0% | 37.5% | 8.3% | 8.3% |
| BnH | 18.3% | 11.7% | 21.7% | 40.0% | 1.7% | 18.3% | 20.0% | 5.0% |

| | | $X^2=3.0$ | $X^2=1.7$ | | $X^2=5.7$ | $X^2=3.5$ | |
| | n (N-item 1) | N.S. | N.S. | | N.S. | N.S. | |

| | |
|---|---|
| Drammen | 38 |
| Munich | 19 |
| Hamburg | 24 |
| Benton Harbor | 60 |

Note:  Items will not add to 100%, because not mutually exclusive.

Table H-13. Herbert's Specified Feelings Regarding Axe
(Category 33)

a. Presence or absence of specified feelings

|  | N | Item(s) 1 None spec. | Item(s) 2-6 Feelings spec. |
|---|---|---|---|
| Drammen | 101 | 56.4% | 43.6% |
| Munich | 49 | 71.5% | 28.5% |
| Hamburg | 50 | 62.0% | 38.0% |
| Benton Harbor | 99 | 50.5% | 49.5% |

$X^2 = 6.3$   Not Significant $(.05 < P < .10)$

b. Kinds of feelings specified -- angry vs. sad

|  | n ($\Sigma$ items 2-4) | Item(s) 2,7 Angry, annoyed | Item(s) 4 Sad |
|---|---|---|---|
| Drammen | 41 | 82.9% | 17.1% |
| Munich | 13 | 53.8% | 46.2% |
| Hamburg | 16 | 50.0% | 50.0% |
| Benton Harbor | 46 | 58.7% | 41.3% |

$X^2 = 8.9$   $P < .05$

Note: Of the "angry, annoyed" group, almost all specified anger. Items 5 ("angry and sad") and 6 ("other") appeared rarely.

Table H-14. Billy's Specified Feelings (Category 35)

|  | N | Item(s) 1 None spec. | Item(s) 2-8 Feelings spec. |
|---|---|---|---|
| Drammen | 101 | 90.1% | 9.9% |
| Munich | 49 | 85.7% | 14.3% |
| Hamburg | 50 | 94.0% | 6.0% |
| Benton Harbor | 99 | 91.9% | 8.1% |

$X^2 = 2.3$   Not Significant

Table H-15.  Restoration of the Axe (Category 37)

| | N | Item 1 Not restored | Item 2 New axe | Item 3 Repaired axe |
|---|---|---|---|---|
| Drammen | 101 | 58.4% | 17.8% | 23.8% |
| Munich | 49 | 73.4% | 14.3% | 12.2% |
| Hamburg | 50 | 48.0% | 30.0% | 22.0% |
| Benton Harbor | 99 | 64.6% | 25.3% | 10.1% |

$$X^2 = 13.9 \qquad P < .01$$