This report is a working paper in the HSR small group research program. Ongoing research in that program, supported by AFOSR under Contract AF49(638)256, has as its objective the development and application of a systematic framework for integration of existing research knowledge of small groups. The research program was initiated with a pilot study, which is reported in (2). The major end product of the pilot study was the development of a classification system which categorized small group studies into nine categories, on the basis of the form or syntax of research information which they contained rather than on the
20. basis of content. Further steps in the pilot study consisted of detailed classification of research information in a sample of studies of one of the syntactical categories—studies of small group effectiveness. The purpose of the present research is to extend the two classification systems developed earlier into one integrated framework for classification and comparison of all types of small group information. This report presents a tentative version of such a model or framework for review. When the present version of the model has been reviewed and pre-tested, it will be used as the basis for a data-recording system which will permit computer processing of research information to be compiled from a comprehensive sample of studies.
A MODEL FOR INTEGRATION OF SMALL GROUP RESEARCH INFORMATION

A Working Paper in the HSR Small Group Research Program

Joseph F. McGrath
Irwin Altman

May 1958

HSR Research Memorandum: 58-1

HUMAN SCIENCES RESEARCH, INC.
1408 North Fillmore Street
Arlington 10, Virginia

Approved for public release; distribution unlimited.
INTRODUCTION

Background

This report is a working paper in the HSR small group research program. Ongoing research in that program, supported by AFOSR under Contract AF 49(638)-256, has as its objective the development and application of a systematic framework for integration of existing research knowledge of small groups.

The research program was initiated with a pilot study, which is reported in (2). One major end product of the pilot study was the development of a classification system which categorized small group studies into nine categories, on the basis of the form or syntax of research information which they contained rather than on the basis of content. Further steps in the pilot study consisted of detailed classification of research information in a sample of studies of one of the syntactical categories—studies of small group effectiveness.

The purpose of the present research is to extend the two classification systems developed earlier into one integrated framework for classification and comparison of all types of small group information. This report presents a tentative version of such a model or framework for review. When the present version of the model has been reviewed and pre-tested, it will be used as the basis for a data-recording system which will permit computer processing of research information to be compiled from a comprehensive sample of studies.
Basic Rationale

The model presented below is designed to provide an integrated framework within which research information from a wide range of small group research studies can be classified, compiled and integrated. The model is formulated in terms of three levels of discourse: relationships, variables and data. Certain distinctions, considered to be fundamental parameters, are delineated at each of these levels of discussion, and appropriate categories or levels of each of these parameters are specified. Relationships among the several parameters at each level of discourse are described. Finally, rules for mapping from one level of discourse to the next are specified.

A relationship is a specific test of the covariation between a pair of variables. Operationally, both of the variables and the test of covariation between them are represented by indices. Such a relationship is considered the unit of research information. See Part A.

A variable is a conceptual abstraction from reality which serves as a focus of study. It is always represented operationally by one or more indices. Classification of a variable is determined by the characteristics of the data upon which the index is based, and the indexing operations used in combining these data. See Part B.

Data are recorded abstractions from reality, which can be described in terms of six basic parameters. See Part C.
Individual data are combined, by one of several indexing operations, to form an abstraction at a higher level which can be called an index. An index is an operational representation of a variable. See Part D.

This general framework is diagramed on page 4.
THE DEVELOPMENT OF RESEARCH RELATIONSHIPS

Part of Report

PART A

PART B

PART D

PART C

Level of Discourse

Part of Report

Relationship

Variables

Index Levels

Indexing Operation

Data Items

Reality

The Phenomenon of Study

Part of Report

Level of Discourse

PART A

PART B

PART D

PART C

The Phenomenon of Study

Part of Report

Level of Discourse

PART A

PART B

PART D

PART C

The Phenomenon of Study

Part of Report

Level of Discourse

PART A

PART B

PART D

PART C

The Phenomenon of Study
PART A  SYNTACTICAL FORMS OF RELATIONSHIPS

The basic unit of research information is defined as a statement of the result of a specific empirical test of covariation between a pair of variables. Generically, such a relationship can be expressed as \( Y = (f)X \), or \( Y \) varies in some predictable fashion as a function of variation in \( X \).

The term "\( Y \)" will be referred to as the Focal variable, or more specifically as an index of the Focal variable for the stated relationship. By Focal variable is meant the phenomenon or aspect of the phenomenon which is to be accounted for. Similarly, the term "\( X \)" will be referred to as (an index of) the Source variable, meaning that \( X \) stands for a phenomenon or condition which is presumed to partially account for variation in the Focal phenomenon. The term "(f)" will be referred to as the Relational Term, which expresses the direction, functional form and/or degree of variation in \( Y \) which is predictable from known variation in \( X \).

The Relational Term and the variables which the term relates can be classified along a number of descriptive parameters. This paper is concerned with elaborating a classification system applicable to variables.

All relationships are assumed to be directional - variation in the source variable is interpreted as accounting for variation in the Focal variable. Several rules may be applied for determining the directionality of a given empirical relationship.
Temporal referent: The Source variable of a given relationship must be temporally simultaneous with or prior to the Focal variable.

Experimental operations: An experimentally manipulated variable is necessarily Source variable.

Interdependence: When neither a nor b applies, an empirical relationship will be treated as a symmetrical pair of relationships unless the investigator assumes a directionality. When this rule is applied, subsequent information may modify the tentative assignment by indicating that the relationship holds in only one direction or by providing evidence that the relationship does in fact hold in either direction.

The designation of a variable as Source or Focal variable applies only for the specified empirical relationship. The terms Source and Focal are not a classification of the variable per se, but refer to the functional use of the variable in a given relationship. Hypothetically, any variable can be used as either Source or Focal variable in different empirical relationships, though the variable can only serve in one or the other functional capacity within any one test of a relationship.
PART B: CLASSIFICATION OF VARIABLES

In Part A we defined the unit relationship and discussed the functional use of variables within such relationships. This section is concerned with definition and classification of the variables. Subsequent parts of the report will describe features of data and indexing procedures, and will describe how data can be combined to form operational indices of each of the variable classes.

1. Bases of Classification

Three formal parameters of variables are described below:

a. Nature of the Phenomenon

A variable can refer either to a static phenomenon (entity) or to a dynamic phenomenon (event). A variable which has a static referent (STATE) provides a summary of a particular aspect of the entity up to a specified point in time. Thus, it is cumulative in time. A variable which has a dynamic referent (ACTION) provides a description of an event during an interval of time. Thus, it is non-cumulative in time. State precedes and gives rise to Action, and Action may produce subsequent changes in State. Thus, State and Action provide a continuous alternation over time.

b. Task, or Terms of Abstraction

A variable may be abstracted from reality in descriptive or evaluative terms. Description implies specification of a property of the phenomenon of concern. Evaluation implies comparison of the phenomenon to some standard with respect to a property. The property may be continuous or categorical, global or specific. The standard of comparison may be subjective or objective. EVALUATION implies prior DESCRIPTION and specification of a standard on the property.

c. Level of Reference

A variable may refer to a part or to the whole of the unit of observation, or to some aspect of the environment which surrounds that unit. For small group research, the unit of observation is assumed to be a multi-member group. Variables may refer to a part (MEMBER), to the whole...
unit (GROUP), or to the environment (SURROUND). Surround variables include those which refer to individuals and groups which are not a part of the referent group, as well as non-human aspects of the member’s or group’s environment. The levels of reference are related in terms of inclusion. A member is considered embedded in both the group and the surround. The group is embedded in its surround. Variables will be classified at the most restricted level to which they can logically be assigned. Group variables are those which are not attributable to members; surround variables are those which are not attributable to either members or the group as a whole.

These three classifications are independent, in that a variable described by a particular category on one, does not restrict what categories can hold for the other two classifications. Together they generate twelve major classes of variables which repeat along one axis over time, as shown in the chart below.

**CLASSES OF VARIABLES**

<table>
<thead>
<tr>
<th>TASK</th>
<th>LEVEL</th>
<th>NATURE*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>I. State</td>
</tr>
<tr>
<td>A. Description</td>
<td>Member</td>
<td>IAm</td>
</tr>
<tr>
<td></td>
<td>Group</td>
<td>IAg</td>
</tr>
<tr>
<td></td>
<td>Surround</td>
<td>IAs</td>
</tr>
<tr>
<td>B. Evaluation</td>
<td>Member</td>
<td>IBm</td>
</tr>
<tr>
<td></td>
<td>Group</td>
<td>IBg</td>
</tr>
<tr>
<td></td>
<td>Surround</td>
<td>IBs</td>
</tr>
</tbody>
</table>

*State-Action alternate over time.
2. Span of a Relationship

The span of a relationship is defined as the number of variable class boundaries which separate the Source and Focal variable of that relationship. One class boundary separates State from Action, another boundary separates Action from State at a subsequent time. One boundary separates Description from Evaluation. One boundary separates Member from Group; one boundary separates Group from Environment; one boundary separates Member from Surround (although certain effects of Surround may be mediated via the Group to the Member and vice versa.

Thus, for example: A relationship between a pair of variables both of which describe some aspect of Member State (cell IAm) has a span of zero; a relationship between a description of Member State (cell IAm) and a description of Member Action (cell IIAm) has a span of one. A relationship between a description of Member State (cell IAm) and an evaluation of Group Action (cell IIBg) has a span of three, since the Source and Focal variables are separated by the State-Action boundary, the Description-Evaluation boundary, and the Member-Group boundary.

The span of a relationship is a rough index of the degree of probable mediation which underlies that relationship, and indicates the variable classes which are most appropriate to investigate as possible mediating variables. For example: A relationship with zero span, which has Source and Focal variable from the same variable class, can only be mediated by some other variable from that class. A span one relationship, which has Source and
Focal variables from adjacent classes, can be mediated by another variable from either of those classes. A span two relationship, in which the Source and Focal variables are separated by a boundary on each of two classifications, can be mediated by variables from any of four classes.

Thus, when the Source and Focal variables of a relationship have been classified, the framework provided by the variable classification helps to guide the search for mediators, antecedents, and consequences of that relationship.
PART C: CLASSIFICATION OF DATA

The previous parts of this report have described classifications at the level of relationships and variables respectively. This part provides a classification at the level of the individual data item, which is the basic abstraction from reality (See chart page 4). Part D describes procedures by which basic data items are combined into indices that serve as operational definitions of the variable classes discussed in Part B.

An individual datum is defined as the specification of some property of a phenomenon, by an agent, from a particular viewpoint or standard. A datum can be characterized in terms of six basic parameters:

1. Nature of the phenomenon
2. Judgment task
3. Object
4. Agent
5. Standard
6. Item or property

Definitions and categories of these six data parameters are discussed below, and certain contingencies among these parameters are specified. A chart of classes of data generated by combinations of categories on the six parameters is presented on page 16.
1. **Nature of the Phenomenon**

The phenomenon to which a datum refers may be an entity (STATE) or an event (ACTION). There is always a one-to-one mapping from data level to variable level with respect to Nature of the Phenomenon; hence, the same parameter and the same categories appear in both the data and the variable classifications.

2. **Judgment Task**

A datum may consist of the specification of (amount or degree) a property possessed by the phenomenon (DESCRIPTION); or it may state the distance or discrepancy between the phenomenon and some standard with respect to a property (EVALUATION). As with Nature, there is always a one-to-one mapping between data level and variable level with respect to the Judgment Task; hence, this parameter appears in both data and variable classifications.

3. **Object**

A datum may have reference to the unit of study (GROUP), a part of that unit (MEMBER), or some aspect of the medium within which the unit is situated (SURROUND). Note that there is not a one-to-one mapping between the object of a datum and the level of reference of variables which are built up from combination of that and other data. (See Part D for further discussion).
4. Agent

A datum is recorded by some person or by a data-recording device used by some person, i.e., an agent. In small group research a major distinction is needed between data for which some member of the group is data-collection agent (MEMBER) and data for which there is a data-collector who is not a part of the group (INVESTIGATOR). The agent combines with the object of data to help determine the level of reference of variables built from those data.

5. Standard

An evaluative datum may provide a comparison of the phenomenon from the point of view of a member of the group (Subjective) or from the point of view of an external referent (External). A descriptive datum does not involve a comparison to a standard.

6. Item or Property

A datum involves the specification of some content property of the phenomenon. Each data item potentially specifies a different property, although a group of similar items are often used to provide a more reliable estimate of the property. Since there are hypothetically an infinite number of possible properties of a given object, no specific categories of items (properties) are used in the general model. Content properties will provide useful sub-classifications within each of the classes of variables, specific to the class.
There are certain contingencies between these data parameters, in the sense that the occurrence of a given category on one parameter affects which categories can occur on some of the other parameters.

For a Descriptive task, the agent may be either Member or Investigator. When the task is Evaluative, the member may both describe and evaluate the phenomenon, the investigator may do both, or the member may describe and the investigator evaluate.

If the agent is a Member and the object is a Member, a distinction must be made between self and some other member as object. If the agent is Investigator, no self/other distinction is made.

In a descriptive task the phenomenon is not compared to a standard. In an Evaluative task, if the agent is a Member he may compare the phenomenon to his own (Subjective) standard; he may compare it to the (Projective) standard of another group member or the whole group, or he may compare it to an External standard. If the agent is the Investigator, he compares the phenomenon to an External standard.

As indicated previously, item or content classes do not form a general dimension of the data framework, but rather provide a sub-classification within each of the data classes which is specific to that class. For example, an attitude item refers to State not Action, to Evaluation rather than Description; it makes use of a subjective or projective standard, and the agent is a Member rather than the Investigator. The attitude object may be from Member, Group, or Surround classes.
The data collection agent can be either the Investigator or a Member. When the agent is Investigator, it is usually assumed that the resulting data refer to the data object, not to the agent. However, when the agent is a member, resulting data can be treated either as referring to the object or as providing information about the agent. (For example: Responses by a subject to a Rohrschach card could be considered a description of that object; normally, however, those responses are considered as a description of certain aspects of the subject's "state".)

When data provided by a member-agent are taken as providing information about the object, the distinction between Member and Investigator as agent does not affect the referent of the data. When data from a member-agent are taken as providing information about the agent, in effect the agent becomes data object. That is, he is providing information about himself, regardless of whether the stimulus object he is attending to is a member, group, or surround object. Thus, for example, a subject's response to an attitude item is usually considered as information about the subject rather than as information about the object specified in the item. Consequently, such a datum can be classified as: Agent-Member; Object-Member (self).

The categories of these six data parameters and the limitations between parameters together generate 46 data classes which are shown in the chart on page 16. These data classes do not include item or property sub-classes, which are potentially infinite in number and specific to each of the data classes specified in the chart.
### Classes of Data

<table>
<thead>
<tr>
<th>TASK</th>
<th>AGENT</th>
<th>STANDARD</th>
<th>OBJECT</th>
<th>NATURE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Describe</td>
<td>Member</td>
<td>None</td>
<td>Member</td>
<td>Group</td>
</tr>
<tr>
<td></td>
<td>Investigator</td>
<td>None</td>
<td>Self</td>
<td>Other</td>
</tr>
<tr>
<td>Evaluate</td>
<td>Describer</td>
<td>Member</td>
<td>Subjective</td>
<td>Projective</td>
</tr>
<tr>
<td></td>
<td>Evaluator</td>
<td>Member</td>
<td>Self</td>
<td>Other</td>
</tr>
<tr>
<td></td>
<td>Member</td>
<td>Investigator</td>
<td>External</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Investigator</td>
<td>Investigator</td>
<td>External</td>
<td></td>
</tr>
</tbody>
</table>

16
PART D. INDEXING OPERATIONS AND TRANSFORMATIONS
FROM DATA TO VARIABLE LEVEL

Previous sections have described classifications at relationship, variable, and data levels, respectively. This part of the paper deals with indexing operations which, together with the data classification, provides rules for mapping from the data level to the variable level.

1. Indexing Operations

An index is formed by combining a number of data with respect to one or more of the data parameters. The combination of data with respect to a single parameter is an indexing operation. Successive indexing operations may be used to transform data with respect to two or more data parameters.

Basically, data can be combined in either of two ways: by summation, to provide an estimate of total, or of central tendency; or by subtraction, to provide an estimate of scatter or discrepancy. The term \( \sum \) will be used to designate the former, and the term \( \triangle \) will be used to designate the latter.¹

When data are summed with respect to a certain parameter, the resulting index refers to a characteristic of the individual datum upon which the summing operation was carried out. When data are combined by some discrepancy the resulting index refers to a characteristic of the combination of those individual data. For example:

¹¹ The term \( \sum \) refers to a number of additive operations including summation, covering, frequency counts, etc. The term \( \triangle \) refers to a number of discrepancy operations including variance, standard deviation, profile scores, etc.
a. If the investigator adds a single member's responses to a number of similar items, the resulting "scale score" is an estimate of the difficulty of the average item in the scale. If the investigator computes a variance among the member's responses, the resulting index relates to the heterogeneity of the scale items, i.e., the index refers to a relationship among items rather than to items per se.

b. If the investigator sums the responses of many members to a single item, the resulting index is an estimate of the average member response to the item. If the investigator computes a variance among the members' responses to a single item, the index refers to the heterogeneity of member responses and thus has a group referent, i.e., the index refers to a relationship among members rather than to members per se.

Thus, the referent of the resulting index is partly determined by whether the data were combined by a summing or a discrepancy procedure. This distinction plays an important part in determining the mappings from data classes to classes of variables.

Mapping from Data Classes to Variable Classes

Three bases for classification of variables were described in Part B: Nature (State or Action), Task (Description or Evaluation); and Level of Reference (Member, Group, or Surround). There is a direct or one-to-one mapping from data to variable class with respect to Nature and Task. That is, as a rule, State and Action items are not combined within the same index, nor are Descriptive and Evaluative items combined within the same index. However, there is not a one-to-one mapping between data and variable classes with respect to the Level of Reference of the variable.
The Level of Reference of the variable classification is determined by the data Object and the type of indexing operation (≤ or △) by which data are combined with respect to Object. For Group or Surround data objects, there is a direct mapping to Group or Surround Level of Reference, respectively, regardless of indexing operations. For Member data objects, if no combination over members is made, or if an additive combination over members is made, the resultant index is at Member Level of Reference. If a discrepancy operation is used to combine over member objects, the resultant index is at Group Level of Reference.

RULES FOR MAPPING FROM DATA CLASS TO LEVEL OF REFERENCE OF VARIABLES

<table>
<thead>
<tr>
<th>Data Object</th>
<th>Indexing Operations</th>
<th>Level of Reference of Variable Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEMBER</td>
<td>None or ≤</td>
<td>Member</td>
</tr>
<tr>
<td></td>
<td>△</td>
<td>Group</td>
</tr>
<tr>
<td>GROUP</td>
<td>No combination can be made</td>
<td>Group</td>
</tr>
<tr>
<td>SURROUND</td>
<td>None, ≤, or △</td>
<td>Surround</td>
</tr>
</tbody>
</table>
Note that these rules for mapping from data to variables do not deal with either Standard or Item parameters. These parameters do not affect the mappings from data to the three bases of classification of variables dealt with in the present formulation of the model. (See Part B) It is likely that additional bases of classification of variables will be necessary, and that mappings from the data level for these will involve the Standard and Item parameters.
PART E. APPLICATION OF THE MODEL AS A FRAMEWORK
FOR CLASSIFYING AND INTEGRATING SMALL
GROUP RESEARCH INFORMATION

The definitions and assumptions stated in Parts A and B imply a matrix
of syntactical forms of unit relationships shown in the chart on page 23. This
matrix has Source and Focal variables as major axes, with Nature, Task and
Level of Reference along each axis, and extended in time along the Focal
variable axis.

Previous definitions and assumptions assign properties and place restric-
tions upon relationships in various cells of the matrix. For example: Cells
in the lower left portion of the matrix are null cells since they specify relation-
ships between Focal variables logically or temporally prior to the Source
variable. The span of the relationship between Source and Focal variables
is indicated by the entries in each cell.

For each relationship investigated in the small group research studies
to be reviewed, the Source and Focal variable will be identified. Each of these
variables will be classified as to Nature, Task, and Level of Reference combi-
nation by examination of the data parameters and indexing operations on which
the variable is based. The relational term will also be classified in terms of
the operations employed to test the relationship between Source and Focal
variables, and the direction and degree of relationship which was obtained.

When each variable has been classified it will be assigned the same
position on both Source and Focal axes, and will be assigned to both initial
and later time periods on the Focal axis. These classification procedures will transpose the matrix of syntactical forms into a master table of possible relationships between pairs of variables. Each axis of the master table will contain the same list of variables in the same order. Within limits of the sample of studies covered, the master table will include all variables generated in small group studies and will contain cells which specify all possible relationships between pairs of these variables.

Each cell of the master table will specify a potential relationship between two specific variables. Entries in cells will describe the direction/degree of the obtained relationship. Duplicate entries in a given cell will represent replicated relationships, while entries in adjacent and nearly adjacent cells will represent closely similar relationships.

Placement of each relationship within this syntactical framework will help guide the search for its mediators, antecedents and consequences. Examination of sets of adjacent cells for example, may reveal chains of relationships of the form: A-B; B-C, C-D, etc. Thus, the model not only provides an integrated system for classification of research information, but also provides a framework within which this information can be synthesized into an integrated body of knowledge, and will point up potential relationships that have not as yet been explored.
# The Model as a Framework for Integrating Research Information

<table>
<thead>
<tr>
<th>Source Variable</th>
<th>Focal Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Time One</td>
</tr>
<tr>
<td>IAm</td>
<td>IAg</td>
</tr>
<tr>
<td>IAm</td>
<td>0</td>
</tr>
<tr>
<td>IAg</td>
<td>1</td>
</tr>
<tr>
<td>IAs</td>
<td>1</td>
</tr>
<tr>
<td>IBm</td>
<td>0</td>
</tr>
<tr>
<td>IBg</td>
<td>1</td>
</tr>
<tr>
<td>IEs</td>
<td>1</td>
</tr>
<tr>
<td>IIAm</td>
<td>0</td>
</tr>
<tr>
<td>IIAg</td>
<td>1</td>
</tr>
<tr>
<td>IIAs</td>
<td>1</td>
</tr>
</tbody>
</table>

Key: I (State); II (Action); Nature of the Phenomenon
A (Description); B (Evaluation); Judgment Task
m (Member); g (Group); s (Surround): Level of Reference

Cell entries indicate Span of the relationship defined by the row and column.

23