INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again–beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

University
Microfilms
International

300 N. ZEEB ROAD, ANN ARBOR, MI 48106
18 BEDFORD ROW, LONDON WC1R 4EJ, ENGLAND

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/4724

8018927

McClanahan, Walter Val

A META-ANALYSIS OF LOCUS OF CONTROL RESEARCH APPEARING IN THE JOURNALS OF THE AMERICAN PSYCHOLOGICAL ASSOCIATION FROM 1963 THROUGH 1977

The University of Oklahoma Ph.D. 1980

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106 18 Bedford Row, London WC1R 4EJ, England

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark ✓.

1. Glossy photographs _____
2. Colored illustrations _____
3. Photographs with dark background _____
4. Illustrations are poor copy _____
5. Print shows through as there is text on both sides of page _____
6. Indistinct, broken or small print on several pages ✓ throughout _____
7. Tightly bound copy with print lost in spine _____
8. Computer printout pages with indistinct print _____
9. Page(s) _____ lacking when material received, and not available from school or author _____
10. Page(s) _____ seem to be missing in numbering only as text follows _____
11. Poor carbon copy _____
12. Not original copy, several pages with blurred type _____
13. Appendix pages are poor copy _____
14. Original copy with light type _____
15. Curling and wrinkled pages _____
16. Other _____

University Microfilms International

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

A META-ANALYSIS OF LOCUS OF CONTROL RESEARCH APPEARING IN THE JOURNALS OF THE AMERICAN PSYCHOLOGICAL ASSOCIATION FROM 1963 THROUGH 1977

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

BY

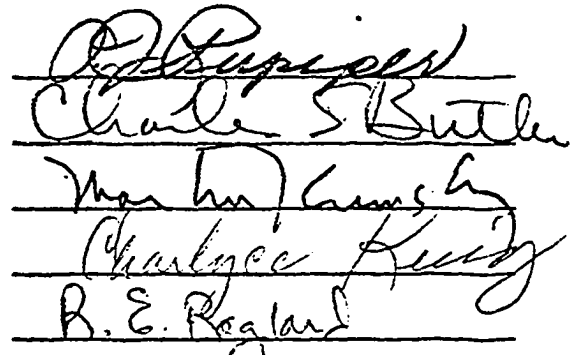
WALTER VAL McCLANAHAN

Norman, Oklahoma

1980

A META-ANALYSIS OF LOCUS OF CONTROL RESEARCH APPEARING IN THE JOURNALS OF THE AMERICAN PSYCHOLOGICAL ASSOCIATION FROM 1963 THROUGH 1977

APPROVED BY



DISSERTATION COMMITTEE

ACKNOWLEDGMENTS

The writer is greatly indebted to Dr. Omer J. Rupiper, the chairman of his doctoral committee, for the invaluable direction, assistance, and support he provided throughout the various stages of this endeavor, from the initial formulation of the problem through the completion of the final manuscript.

The writer wishes to express his sincere appreciation to Dr. Charylce King, Dr. Martin Krimsky, and Dr. Robert Ragland for the constructive assistance and encouragement they provided during the course of his doctoral program, and to Dr. Charles Butler for the assistance he provided as the reading member of the doctoral advisory committee.

In addition, the writer is grateful to Dr. Kit Farwell, his supervisor at the Phil Smalley Children's Center, for the inspiration and support he provided, and to Dr. Gwen Garner and the other officials of the Oklahoma Department of Mental Health for granting him the educational leave to complete this project.

The writer is also indebted to George Calhoun, Michael Daves, Dr. Helen Klein, Sally LeCrone, Dr. Laurel Von Horn, and Jerry Witten, the psychologists who so kindly assisted him in the reliability procedures undertaken in the study.

Finally, the writer wishes to express his deep appreciation to his wife, Cheryl, and his daughter, Valerie,

for their constant encouragement, understanding, patience, and support, and for sacrificing so much, with so little complaint, during the long years of graduate study.

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS	iii
LIST OF TABLES	vi
Chapter	
I. INTRODUCTION	1
Statement of the Problem	16
Definition of Terms	17
II. METHOD	19
Sample	19
Procedure	33
III. RESULTS	39
IV. SUMMARY, CONCLUSIONS, DISCUSSION, AND IMPLICATIONS	75
REFERENCES	86
APPENDICES	
A. List of Sample Articles, with Numerical-Alphabetical Reference Codings, for Sample Studies	92
B. Quality of Study Rating Form and Form for the Evaluation of an Article	104
C. Tables	107
D. Items and Scoring Criteria for Quality of Study Rating Form	124
E. Content Categories and Assignment Criteria	133

LIST OF TABLES

Table		Page
1.	Positive versus Negative Findings by Quality of Study Ratings	41
2.	Positive versus Negative Findings by Source-of-Derivation-of-Hypotheses, Contrast of Positive versus Negative Findings for Hypotheses with LC-Theory versus Previous-Research-Alone Derivations	44
3.	Positive versus Negative Findings by LC Scale Used	48
4.	Positive versus Negative Findings by General-Design-of-Study, Positive versus Negative Findings by True Experimental versus Quasi-Experimental Design of Study	51
5.	Positive versus Negative Findings by Internal-versus-External Determination	54
A.	W Coefficients for Rankings of Item-by-Item Scoring Criteria, Reliability of Scoring Criteria Rankings	108
B.	W Coefficients for Rankings of Item-by-Item Scoring Criteria for the Eight Altered Items, Reliability of Scoring Criteria Rankings for the Eight Altered Items	112
C.	Overall Quality of Study Ratings (Reliability Sample), Reliability of Overall Ratings	114
D.	Assignments to Content Categories, Reliability of Assignments	115
E.	Overall Ratings, Mean Ratings, and Category Placements by Number of Study	116

A META-ANALYSIS OF LOCUS OF CONTROL RESEARCH APPEARING IN THE JOURNALS OF THE AMERICAN PSYCHOLOGICAL ASSOCIATION FROM 1963 THROUGH 1977

CHAPTER I

INTRODUCTION

The past two decades have witnessed a tremendous growth in scientific research and the publication of research findings. Areas of investigation represented by a few dozen references in the 1940s or 1950s now encompass hundreds or even thousands of studies (Glass, 1977), yet all indications suggest that, not only will the trend continue, more than likely it will accelerate. Many researchers have, for some time, found it difficult to keep up with current developments, and have, as noted by Glass and Norward (1958) more than twenty years ago, tended to rely upon chance and hearsay to learn of work pertinent to their own research interests, but the problem, unfortunately, was sure to worsen, as it has been predicted that researchers of the year 2000 will need to read four times as many papers as today simply to maintain current coverage proportions (Kulik, Kulik, and Cohen, 1979).

One potential solution to this problem, first proposed by Glass (1976), was the integration of large collections of related research findings by means of what Glass has termed the meta-analysis of research. In challenging his readers with Eliot's question, "Where is the knowledge we have lost in information?", Glass called for a reordering of priorities in educational research, with a view toward concentrating more scholarly effort on the problem of finding the knowledge that lies untapped in the ever-growing mountain of individual research studies.

Methods of integration advocated by Glass stand out in sharp contrast to those of the traditional narrative review of research: According to Glass,

> A common method of integrating several studies with conflicting findings is to carp on the design or analysis deficiencies of all but a few studies--those remaining frequently being one's own work or that of one's students or friends--and then advance the one or two "acceptable" studies as the truth of the matter (Glass, 1976, p. 4).

Methods of integration used in meta-analysis, by contrast, consisted of "the same statistical methods applied in primary data analysis" (Glass, 1976, p. 6).

While Glass's work was seminal to the field of meta-analysis research two limitations of his methods may be identified. The first had to do with the boundaries of meta-analysis, or rather, the population of research topics considered appropriate to meta-analytic study. While Glass (1977) argued against the position that only studies which

were the same in certain respects could be aggregated, his own meta-analytic research efforts, and those of all to date who have followed his lead, have been confined to analyses of collections of studies which were homogeneous with respect to the major variables investigated. Thus, White (1976), in the very first meta-analysis, investigated, through an analysis of 636 correlational findings, the relationship between socioeconomic level and academic achievement; Smith and Glass (1977) investigated the "effect size" of 400 studies of the effects of psychotherapy; and Kulik, Kulik, and Cohen (1979) examined the findings of 75 studies concerning the effects of Keller's Personalized System of Instruction on academic performance, course completions, course ratings, and student study time. While there is no dearth, at this point in time, of collections of homogeneous studies to be analyzed, the meta-analysis concept would seem to lend itself to wider applications, i.e., to collections of studies more heterogeneous with respect to variables investigated, as suggested by some of the efforts that have been undertaken in the field of program evaluation in recent years (Cook and Gruder, 1978).

The second limitation of Glass's approach was concerned with the relative weight assigned to the findings of studies varying with respect to sampling, data-collection, design, and analysis considerations. Here, Glass warned against taking design and analysis too seriously, pointing

out that "a study with a half-dozen design and analysis flaws may still be valid" (Glass, 1976, p. 4). He advocated including in meta-analysis all studies which meet very basic, minimal design and analysis standards, and assigning equal weight to the findings of all studies included, noting that his own past experience had led him to question "whether well-designed and poorly-designed experiments give very different findings" (p. 4). In his previously-cited 1977 article, Glass continued to develop a rather compelling argument for the inclusion of flawed studies in the meta-analysis, pointing out that a set of consistent findings from a group of weak studies could still add up to a strong conclusion, i.e., if the locus of deficiencies varied across studies, an appeal to parsimony would favor the simple explanation of validity of findings rather than the multiple competing explanations necessary to account for such a set of consistently spurious findings.

Glass's point was well taken, since potentially valuable information could be obtained from even flawed research studies. It was the present investigator's opinion, however, that Glass had taken this one step too far, since little to no confidence could be placed in the findings of a seriously flawed study, and these must not be permitted to obscure, or reverse, the direction of the results of much more adequate efforts. Mansfield and Busse addressed this issue in their "rejoinder to Glass" (1977), noting

that their experience in summarizing the research on creativity indicated to them that well-designed and poorly-designed studies do tend to yield quite different results at times:

> . . . well-designed studies are less likely to yield significant training effects than smaller, poorly designed ones . . . (for), poorly designed studies are likely to yield spurious findings precisely because these studies are poorly designed. (p. 3)

Mansfield and Busse went on to indicate a potential bias effect resulting from the publication rates of large well-designed studies and small, poorly-designed ones, pointing out that the larger, better-designed studies were likely to be published whether predicted relationships were borne out or not, while the smaller, more poorly-designed studies were likely to be published only if they obtained significant findings. The result of this bias, of course, would be to spuriously enhance the treatment effects investigated. They also indicated that they found fault with Glass's computing average effect size across studies, as this assumed too much comparability across investigations.

Glass's approach also stood in opposition to much recent research on the adequacy across situations of different techniques for statistical analysis, such as the studies concerning multiple comparison techniques conducted by Field and Armenakis (1974), Hopkins and Chadbourne (1967) and Petrinovich and Hardyck (1969). These studies have consistently indicated the likelihood of erroneous conclu-

sions due to misuse of various statistical techniques, yet Glass's approach would, presumably, include on an equal footing even those studies whose methods of statistical analysis were such as to virtually guarantee a certain number of significant findings in the absence of any effect whatsoever.

While one may take exception to the specifics of Glass's method, however, the concept of and the possibilities for meta-analysis were considered robust, and the needs indisputable, at least, insofar as the social sciences, particularly psychology and education, were concerned, for it was in these areas that research was expanding at its most rapid rate, and in these areas, too, where the definitive experiment was most elusive, and the conflicting findings most in evidence. The problem was not that reviews in these areas had not been undertaken, for they had been; there were several psychology and education journals, in fact, that published nothing but reviews. The problem was that the traditional narrative review was not equal to the task of organizing, analyzing, and synthesizing such vast bodies of information as had now accumulated. More objective, systematic methods were needed, and, at the present point in time, meta-analytic approaches appeared the most promising.

The research literature on internal-external control of reinforcement, more widely known as internal-external

locus of control, was an example of a rapidly expanding research area in need of such systematic review and analysis. The theoretical construct of internal-external control was derived from Rotter's social learning theory (Rotter, 1954).

Rotter's Social Learning Theory

Social learning theory was concerned with how choices were made by individuals from among the range of potential behaviors available to them. An interaction between the individual and the environment was assumed, but, on the person side of the interaction, dispositional factors derived from the individual's experiences in the social environment were considered much more important than biological factors, though the dispositional factors were seen as being originally derived from "the learned consequences of biological or other built-in reinforcements" (Phares, 1975, p. 15). With respect to the environment side of the interaction, subjective interpretations of the significance of stimuli were considered much more important than the objective properties of the stimuli; the individual responded subjectively to his environment on the basis of his specific learning history and experiences. In essence, then, at this most general level, it was an interaction between the dispositions of the individual and specific situational cues in the environment that was proposed. The dispositional factors were assumed to dominate in ambiguously-

structured situations, the situational ones to dominate in clearly-structured situations.

It was also assumed by social learning theory that the personality exerts an organizing and unifying influence upon the individual's experiences. The impact and significance of a new experience in a given situation was, in large part, determined by the nature of the accumulation of previous experiences in that situation, though allowance was made for change through the impact of reasonably strong, repeated, divergent experience. In like manner, novel experiences were influenced by accumulations of related or similar experiences.

Behavior was considered to be goal-directed; individuals strived to obtain or to avoid certain objects or situations in their environment. The reinforcement value of the objects or situations in the environment may be determined by the empirical law of effect--an object or situation was considered to have positive reinforcement value if the behavior of the individual was directed toward attainment of the object or involvement in the situation; an object or situation was considered to have negative reinforcement value if the behavior of the individual was oriented toward the avoidance of the object or situation.

Finally, it was further assumed that "the occurrence of a behavior of an individual is determined not only by the nature or importance of goals or reinforcements,

but also by the person's anticipation or expectancy that these goals will occur" (Rotter, 1954, p. 102). Reinforcement alone did not adequately explain behavior. For a certain behavior to occur in a given situation, the individual must have an expectancy that the behavior will, indeed, result in the attainment of the goal he was working toward. Such expectancies were built up through previous experiences with the given behavior in various situations: successful past experience with the behavior in a given situation led to the expectancy that it would probably work again in that same situation in the future; unsuccessful experience would reduce the expectancy that the behavior would lead to the attainment of the goal, in the situation in question, at least.

To determine which behavior had the strongest potential for occurrence in a given situation, the individual's expectancy, the reinforcement value of the object or situation in the environment, and the psychological situation, or the meaningfulness or relevance of the object or situation, must all be taken into consideration. Rotter had presented a formula for this, which had been expressed in the following terms:

$$BP_{X,S_1,Ra} = f(E_{X,R_a,S_1}, \ \& \ RV_{a,s_1})$$

The formula is read as follows: the potential for behavior X to occur in situation one in relation to reinforcement a

is a function of both the expectancy that reinforcement a will occur subsequent to behavior X in situation one, and of the value of reinforcement a in situation one.

A second class of expectancies, similar to learning sets or higher-level learning skills may also be identified. These were what Rotter has called generalized expectancies. Generalized expectancies developed from a generalization of experiences from the situation with which they were originally associated to a set of other, related or similar situations. These were considered capable of producing effects quite independent of the actual reinforcement involved in a situation, but interacted with situation-specific expectancies to produce the overall expectancy expressed in the previous formula. Rotter had presented a formula for the overall expectancy for a given situation 1, too, and it had been expressed in the following terms (an S_1 subscript had been substituted for the previous X subscript, as here we were concerned only with the expectancy, not behavior X):

$$E_{S_1} = \frac{f(E' \ \& \ GE_r \ \& \ GE_{ps_1} \ \& \ GE_{ps_2} \ \ldots \ \& \ GE_{ps_n})}{f(N_{S_1})}$$

The formula is read as follows: a given expectancy in situation one is a function of the specific expectancy that a given reinforcement will occur based upon previous experience in the same situation (E'), experiences generalized from other related situations (GE_r), and a variety of

problem-solving generalized expectancies (GE_{ps_1} ... GE_{ps_n}, generalized expectancies oriented toward the solution of certain classes of problems), divided by some function of the number of experiences the individual has had with the specific situation.

One important example of these generalized expectancies was the individual's locus of control orientation, his perception of the extent to which he, or various external forces, control the reinforcements potentially available to him. As may be seen by reference to the above formulae, locus of control was only one of a complex of factors that interact to produce a given behavior in a given situation.

Locus of Control Measurement and Previous Efforts to Integrate Samples of the Locus of Control Literature

The first crude attempt to operationalize and quantify the internal-external locus of control construct was made by Phares (1955), but it was James (1957) who, building upon the earlier work of Phares, developed the first locus of control scale of modern configuration. Since that time a number of other scales had been developed to measure this construct, but the scale that had received the widest application was one developed by the originator of the construct, Rotter himself. Prociuk and Lussier (1975) reported, for example, that the Rotter I-E Scale (Rotter, 1966) had been employed in 69% of the studies included in their content

analysis and bibliography of 1973-1974 locus of control studies. Despite this popularity, there had been increasing criticism of the Rotter Scale in recent years, with some researchers claiming it was a multi-dimensional instrument (Abrahamson, Schulderman, and Schulderman, 1973; Collins, 1974; Gurin, Gurin, Lao, and Beattie, 1969; Schneider and Parsons, 1970), and that scores on the inappropriate factors would wash out or spuriously enhance effects tapped by the appropriate factor for a given application. Others claimed that it was too general for certain applications, and that other, application-specific locus of control scales should be developed, and used in its place (Rohsenow and O'Leary, 1978; Walker, Nast, Chaney, and O'Leary, 1979; Wallston, Wallston, Kaplan, and Maides, 1976).

As more sophisticated measuring instruments became available, and awareness of the activity in this research area increased, the yearly production of locus of control studies accelerated dramatically. Thornhill, Thornhill, and Youngman (1975) for example, identified over 1200 locus of control references for the period 1954 through 1974. Preliminary investigation by the present researcher indicated that a minimum of an additional 1000 locus of control references had accumulated since 1974, for a total of some 2200 studies. A number of attempts had been made to organize, analyze, and synthesize this vast body of data, e.g., four bibliographies (MacDonald, 1972; MacDonald and Davis,

1974; Thornhill, Thornhill, and Youngman, 1975; Throop and MacDonald, 1971), one content analysis and bibliography (Prociuk and Lussier, 1975), and seven reviews (Joe, 1971; Lefcourt, 1966, 1972; Phares, 1973, 1975; Rotter, 1966, 1975) had been published across the past 15 years. Few of these efforts could be accurately described as comprehensive, however, even with respect to their announced temporal and other limitations, and none could be accurately described as systematic analyses of the literature, for the reviews primarily took the form of narrative descriptions and summaries of findings, with little attempt made to objectively evaluate the quality of the studies reviewed or to assess the implications of differential quality for the conclusions that were drawn. None of the authors who reviewed the studies attempted a statistical analysis of findings.

Particularly salient as a limitation of previous efforts at integration in this area of research was the lack of attention given to the quality of study variable, especially with respect to design and statistical analysis considerations, and considerations regarding internal validity. That these considerations were important to the interpretation of the findings of any study, there can be little doubt, e.g., Campbell and Stanley (1963) described internal validity as "the sine qua non . . . the basic minimum without which any experiment is uninterpretable"

(p. 5). Appropriate use of statistical analysis techniques was also considered extremely important, crucial, in fact, to the level of confidence that could be placed in any given set of findings. Field and Armenakis (1974) pointed out that the use of multiple tests of significance may result, in and of itself, in quite inappropriate conclusions being reached, for, as the number of independent statistical tests increases, the probability of rejecting the null hypothesis, the hypothesis of no difference between groups, increases also--by the formula of $1 - (1 - \alpha)^k$, where α = the level of significance and k = the number of significance tests conducted. Thus, the more tests of significance conducted, the more likely it would be that one or more findings would attain significance due to the operation of chance factors alone. That the present study was not merely setting up straw men for the sake of argument may be demonstrated through a casual inspection of Phares (1975), generally considered the most thorough and comprehensive of the various locus of control reviews. Following is a passage from Locus of Control in Personality (Phares, 1975, pp. 68-69):

> We have discussed at length the relatively greater control that internals exert over their environment. Perhaps related to internals' feelings that they can control the environment is the feeling that they can control themselves. For example, Straits and Sechrest (1963) noted that non-smokers were significantly more internal than smokers. James, Woodruff, and Werner (1965) replicated the finding of Straits and Sechrest. In addition, they noted that in the case of male smokers, those who, following the Surgeon General's Report on the Hazards of Smoking, quit smoking and

> did not return within a given period of time were more internal than those who believed the report but did not quit smoking. For females, the result was not significant. It should be noted, however, that Lichtenstein and Keutzer (1967) could not corroborate these relationships between locus of control and smoking.

It should also be noted, but it was not, that James, Woodruff and Werner (1965) calculated 8 *t* tests, all at $\alpha = .05$, in pursuit of their findings, and that Straits and Sechrest (1963) calculated 14 *t* tests, all at $\alpha = .05$ in pursuit of theirs. Applying the Field and Armenakis (1974) formula--a well accepted formula, by the way; Kirk (1968) presented his own version of it in his book Experimental Design: Procedures for the Behavioral Sciences--to the James, et al. and the Straits and Sechrest data revealed that there was, in the case of the former, not a .05 probability of obtaining at least one significant finding due to the operation of chance factors alone, but a probability potentiality as high as .337 of doing so. In the case of the latter, rather than a .05 probability once again, there was, in fact, a probability potentially as high as .512. In other words, instead of there being, in each case, a 1 in 20 chance of one of the alternative hypotheses being confirmed due to the operation of chance factors alone, there was, potentially, in the one case, 1 chance in 3 of such an even occurring, in the other, 1 chance in 2. It was considered significant also that the findings in ques-

tion here were not confirmed in the subsequent study by Lichtenstein and Keutzer (1967).

Statement of the Problem

The limitations and deficiencies reviewed above suggested a need for further analysis of this body of literature. The present study was designed both to fill a need for a comprehensive, systematic review which would take into account the variable quality of the studies published in the locus of control research area, and to explore possibilities for the statistical analysis of a collection of studies which were heterogeneous with respect to variables measured.

The present study proposed to review the locus of control literature from 1957 through 1977 in a systematic manner which would take into account the variability in the quality of published research studies, and to conduct a statistical analysis of the findings. It was not considered necessary to provide coverage prior to the 1957 James Study, as the locus of control variable was not fully operationalized until that time, and the present investigation was concerned only with quantitative data relating to this construct. In addition to the above, the study also attempted to:

1. Investigate the construct validity of the locus of control construct through an examination of findings

within the context of sources-of-derivation-of-hypotheses categories, in addition to the overall analysis.

2. Investigate possible differentials related to locus of control measurement procedures through an examination of findings within the context of measurement-scale-selected categories.

3. Investigate possible differentials related to design of study through an examination of findings within the context of general-design-of-study categories.

4. Investigate possible differentials related to criteria for locus of control group-assignment through an examination of findings within the context of internal-versus-external-determination categories.

Definition of Terms

1. Content analysis--a systematic, objective method of grouping and analyzing communications and communication content to obtain measures of variables.

2. Content categories--exhaustive, mutually-exclusive compartments into which material is grouped for content analysis.

3. External locus of control--external control refers to the perception that the occurrence of an event, positive or negative, is unrelated to one's own actions, and is thus beyond personal control.

4. Internal locus of control--internal control

refers to the perception that the occurrence of an event, positive or negative, is a consequence of one's own actions, and is thus under personal control.

5. Meta-analysis--the analysis of analyses, or, the analysis of the results of a large collection of research studies for the purpose of synthesizing and integrating the findings.

CHAPTER II

METHOD

Sample

Since a systematic review of over 2200 locus of control studies was beyond the resources of the present investigation, analysis was confined to a sample of the population of locus of control studies. Since the journals of the American Psychological Association provided a broader coverage of the various areas and specialties of psychology than those of any other publisher, and since two of the APA journals, the Journal of Consulting and Clinical Psychology and the Journal of Personality and Social Psychology, published rather sizable proportions of the locus of control studies published in journals--only Psychological Reports published more locus of control studies in 1973-1974, for example (Prociuk and Lussier, 1975), than these two--it was decided that all locus of control studies reported in the journals of the American Psychological Association from 1957 (subsequent, that is, to the 1957 James study) through 1977, which met the following set of specific criteria, would be included in the present investigation. The specific criteria for inclusion in the sample were:

1. Journal articles in which studies were reported must be greater than one page in length. The reason for this criterion was that it was felt a one-page report would not be able to provide enough information for a reasonably adequate rating of the study to be obtained.

2. Studies must obtain a measure of locus of control per se, or of internal versus external orientation. As noted in the Purpose of the Study section, the present investigation was concerned only with quantitative data relating to this construct.

3. Studies must test hypotheses, and at least one of these must be concerned with a locus of control issue. There were two reasons for this criterion. The first had to do, once again, with the purpose of the study. One of the purposes of the present investigation was to bring further, hopefully more objective, evidence to bear on the construct validity of the locus of control construct. One of the better ways to accomplish this, it was felt, was to examine proportions of positive-versus-negative findings from studies whose hypotheses were testing relationships derived from social-learning-theory-predictions regarding locus of control. The second reason had to do with sheer practicality; limiting the sample to studies testing hypotheses, and limiting findings examined to those relating to hypotheses concerned with a locus of control issue, both made a formidable undertaking a bit more manageable, and

provided for more objective standards for inclusion of findings for analysis.

The final sample consisted of 150 studies taken from 137 APA journal articles. A complete list of the articles, arranged alphabetically, with numerical-alphabetical reference codings, is provided in Appendix A. The codings from this list will be used to identify the studies in all subsequent tables.

Selection of Studies. Locus of control studies appearing in APA journals from 1971 through 1977 were located by a search of the indexes of Volumes 46 (July-December, 1971) through 61 (January-June, 1979) of Psychological Abstracts. Studies appearing in APA journals from 1957 through 1970 were located by a search of the bibliographies developed by Thornhill, Thornhill, and Youngman (1975), which covered the period 1954 through 1974, and Throop and MacDonald (1971), which covered the period 1954 through 1969, and by inspection of the references of the reviews by Lefcourt (1966) and by Rotter (1966). Both of these reviews and the Throop and MacDonald bibliography have been considered as being quite comprehensive (Joe, 1971; Thornhill, Thornhill, and Youngman, 1975). It was hoped that by this combination of searches all, or nearly all, of the studies would be isolated.

Inspection of the bibliographies and the references of the two reviews resulted in the location of 336 articles

concerning locus of control which appeared in the various APA journals during the period of interest. Of these, 48 were located through the Lefcourt, Rotter, and Throop and MacDonald searches, 131 were located through the Thornhill, Thornhill, and Youngman search, and 157 were located through the Psychological Abstracts search. Many of the articles were revealed in more than one search; they were credited above to the search in which they first appeared.

A number of the articles located were brief reports, articles too short to provide sufficient data for an adequate evaluation. Addresses of the authors of the articles were obtained from a search of the 1978 Directory of the APA and the 1979 APA Membership Register, and 46 requests for extended reports were sent out. Responses were received from 43 of these, for a 93.8% rate of return. Of these, 21 articles were rejected, either because the studies reported in them did not meet all requirements for inclusion in the sample, or because the researchers contacted indicated that extended reports were no longer available. Four articles were rejected because the extended reports on them did not arrive until after the cut-off date (because of time-frame requirements, only eight weeks could be allowed for the extended reports to arrive). Of the studies described in extended reports which did arrive before the eight-week cutoff, 18 met all sample criteria and thus were included in the present study. Of the total 336 articles,

137 reported studies which met all criteria for inclusion in the sample; a total of 150 studies were reported in the sample articles.

Rating Form

Assessment of levels of quality for the various studies in the sample required the location or development of an appropriate rating scale. Several article-rating scales and sets of research evaluation criteria, to include the manuscript rating forms utilized by AERA and APA reviewers, were examined, but Wandt's Form for the Evaluation of an Article, presented in the Handbook in Research and Evaluation (Isaac and Michael, 1971, p. 156) was considered the most appropriate and comprehensive for the present purposes, therefore it was selected as the basis for the study's rating scale.

Wandt's form was developed for the evaluation of an article, not a study per se, and did not address all areas deemed important, by the present researcher, to the evaluation of a study. Accordingly, further review of evaluation criteria sources was undertaken. Included in this review were Galfo's Criteria for Evaluating Research Studies (Galfo, 1975, pp. 241-248), and sections in Understanding Educational Research (Van Dalen, 1973, pp. 409-419) and Research Methods in Education (Wiersma, 1969, pp. 323-329), titled, respectively, "Evaluation and Publication of Research" and

"Meaningful Reading and Critical Analysis (of Research Articles)," among other sources.

Based upon Wandt's Form for the Evaluation of an Article, and a review of the aforementioned evaluation criteria sources, a 16-item Quality of Study Rating Form was developed. Items 11, 12, and 13 of the Quality of Study Rating Form were borrowed directly from Wandt's form (see Appendix B for both the Quality of Study Rating Form and the Form for the Evaluation of an Article); Items 7, 10, 14, and 15 are only slight rewordings or modifications of Wandt's Items 14, 9, 22, and 25, respectively. Items 1 and 4 are rewordings of criteria presented, respectively, in Wiersma (1969, p. 324) and Galfo (1975, p. 244). The remaining seven items were more generally derived from the aforementioned review.

The Quality of Study Rating Form was originally intended to be used without specific scoring criteria, as is the case with Wandt's instrument, but the low reliabilities obtained during a pilot evaluation of the form necessitated the development of specific scoring criteria arranged on a 5-point rating scale for each of the 16 items of the form. Though point-by-point specific scoring criteria were added, Wandt's descriptors for the five levels of quality were retained. Descriptors for ratings of 5 on down to 2 were incorporated without change; the descriptor for a rating of 1 was changed from "Totally Incompetent" to "Totally Unacceptable."

Scoring criteria for Item 6 were derived from a review of Kerlinger's "Designs of Research" section in his book Foundations of Behavioral Research (Kerlinger, 1973, pp. 299-423). The ordering of the criteria was based upon Kerlinger's statement, "In short, a perfectly conducted experimental research, which yields information that A is related to B, is more trustworthy than a perfectly conducted ex post facto research" (Kerlinger, 1973, p. 316). Scoring criteria for Item 9 were based upon a review of Anastasi's chapter on reliability in her book Psychological Testing (Anastasi, 1968, pp. 71-98), and upon observation of claims regarding levels of reliability commonly presented in research articles. Scoring criteria for Item 11 were derived from a review of Campbell and Stanley (1963); ordering of the criteria was based, too, upon Campbell and Stanley, and also, as above, upon Kerlinger's statement comparing experimental and ex post facto research. Scoring criteria for Item 12 were based upon reviews of Field and Armenakis (1974), Guilford (1965), Hopkins and Chadbourn (1967), Kirk (1968), and Petrinovich and Hardyck (1969). Scoring criteria for the remaining items were developed through application of a graduation of wording procedure designed to enhance discrimination of levels through change in criteria-wording and verbal qualifiers utilized, alone. While no specific rules were developed for this procedure, it can, perhaps, be illustrated through the presentation

of one of the scale items and the criteria associated with it. Item 7, with associated scoring criteria, follows:

7. Data-gathering methods or procedures are appropriate for testing the hypotheses of the study.

Scoring Points for the Five Quality of Study Ratings for Item 7:

_5_ The methods or procedures selected appear ideally suited for obtaining the data needed.

_4_ The methods or procedures selected may not be the best available, but they can do a good job of obtaining the data needed.

_3_ The methods or procedures selected are adequate, nothing more.

_2_ It is questionable whether the data needed can be obtained through the methods or procedures selected.

_1_ The data needed cannot be obtained through the methods or procedures selected.

To provide some evidence that the researcher's ordering of the criteria had not been an arbitrary or idiosyncratic matter, and to obtain data regarding the effectiveness and legitimacy of the graduation of wording procedure, scoring criteria for each item were randomly ordered--through utilization of Snedecor's random numbers table, presented in Wyatt and Bridges (1967)--and submitted to three psychologists for rank-ordering. There was, of course, no indication provided as to the order the

researcher had assigned to the various criteria. Instructions provided to the rankers follow:

> A 16-item Quality of Study Rating Scale is currently under development. This instrument attempts to assess, in a reasonably objective manner, a study's status with respect to major development, design, data-collection, analysis, and presentation variables, measured across a 5-point scale. Scoring criteria have been developed for each point of the scale for each individual item. For purposes of obtaining a measure of agreement regarding a proposed rank-ordering of the scoring criteria, scoring criteria are presented herein in a random order following the item to which they pertain. Please read each item and the five criteria following it, and record in the blank immediately to the left of each criteria, the rank, from 1 to 5, which indicates, in your best judgement, its proper place in the 5-point order. Each rank may be used only once per item. Verbal descriptors associated with the five ranks follow:

Level of Quality	Description
5	Excellent
4	Good
3	Mediocre

2 Poor

1 Totally Unacceptable

Included also, herein, is a brief discussion of the Campbell and Stanley (1963) threats to internal and external validity.

Analysis of the results of these rankings, by means of Kendall's coefficient of concordance, W, a measure of the degree of agreement or association between rankings (Kerlinger, 1973, pp. 292-295), resulted in a coefficient of .949 (see Table A, Appendix C, for rankings of criteria for each ranker). The two most experienced rankers disagreed with the researcher's rankings of the criteria for a 3-rating and a 2-rating on Item 9, reversing the 3-2 order, in each case, on the item. It was decided to yield to their judgement in this matter, and the researcher's order was modified, for this item only, to agree with theirs. Further refinement of the scale, undertaken to enhance clarity and improve coverage of scoring situations, resulted in changes of wording for some criteria. Items effected by these were re-submitted to the above procedure, resulting in a coefficient of concordance of .97 for the eight items involved (see Table B, Appendix C for rankings of criteria for each ranker).

During the pilot evaluation of the rating form, the mean of the ratings assigned to the various items was

utilized as a summary measure of overall quality of study. It was found, however, that this procedure resulted in a rather restricted range of overall ratings, from approximately 2 to 4, from a potential range of 1 to 5; therefore it was decided to develop specific criteria for assignment of overall ratings, so that the overall rating could range across the entire 5-point scale too, rather than across only three of its points. To facilitate discriminations between levels, a distinction was drawn between critical and non-critical items. A critical item was defined as one for which a low score, a rating of 1 or 2, would call into question the findings of the study. Items identified as critical were Items 1, 2, 6, 7, 9, 10, 11, and 12. The scoring criterion for assignment of an overall 5 required, at the minimum, that ratings assigned to the more objective critical items (Items 6, 9, and 11) be nothing but 5s, while ratings down to 3 were permitted for the more subjective critical items. The scoring criterion for an overall 4 required, essentially, that the study received a preponderance of 4- and 5-ratings. Overall 3s were reserved for studies which received more ratings of 3 and below than they did of 4 and above, and for those which received a 2-rating, a "poor," on a critical item (and for which, thus, was introduced some question regarding the validity of the study's findings). Studies which received more than one 2-rating on a critical item were assigned an overall 2 by

the scoring criteria for the overall rating, and studies which received a 1-rating on any critical item were to be assigned an overall 1. See Appendix D for item-by-item and overall scoring criteria for the Quality of Study Rating Form.

The Quality of Study Rating Form was then subjected to an evaluation procedure designed to assess the reliability of the overall quality of study ratings. A reliability sample of 25 articles, representing 27 studies, was randomly selected, through utilization, once again, of Snedecor's random numbers table, from the sample of 137 articles included in the present investigation. These 25 articles, along with a copy of the rating form and scoring criteria (referred to as a scoring booklet), were given to a second set of three psychologists, who were asked to rate each of the 27 studies reported in the articles. Instructions provided the raters follow:

> Record the title of the article in which the study to be rated appears, and the name of the researchers who conducted the study, in the spaces provided at the top left of the Quality of Study Rating Form blank. Rate each study on a scale from 1 to 5 (see 5-point rating scale, below), with respect to how well it meets the criteria pertaining to the 16 items included in the Quality of Study Rating Form. A brief scoring booklet has been included

with the present materials. The socring booklet presents point-by-point instructions for the assignment of ratings on each of the 16 items included in the form. You should become thoroughly familiar with it before attempting any ratings. The 5-point rating scale follows:

Level of Quality	Description
5	Excellent
4	Good
3	Mediocre
2	Poor
1	Totally Unacceptable

Note: Where more than one study is reported in an article, rate each study in the article separately, record the ratings on separate rating form blanks, and identify the study by recording, in the upper right-hand corner of the rating form blank, the number by which it had been designated in the article in which it appeared.

Analysis of the overall ratings assigned to the 27 studies in the reliability sample, by means of the reliability coefficient r_{11}, utilized by Wandt (1967) in his reliability procedure for his Form for the Evaluation of an Article, and presented in Lindquist (1953, pp. 357-362), resulted in a reliability coefficient of .732 (see Table C, Appendix C, for ratings assigned by each rater).

Content Categories

Prociuk and Lussier developed a set of 17 content categories for their content analysis of 1973-1974 locus of control studies. As Prociuk and Lussier did not, however, develop explicit criteria for assignment of articles to categories (Prociuk, 1979), it was necessary for the present researcher to develop his own. Criteria for assignment of articles to content categories were derived from an induction of descriptors obtained through a review of the abstracts of all 137 articles included in the study. Studies were assigned to content categories based, however, upon examination of the entire text of the article. This was done concurrent with the rating of the study or studies reported therein. The 17 content categories utilized, along with the assignment criteria pertaining to them, are presented in Appendix E.

The reliability of the assignment procedure was tested by familiarizing one of the three psychologists rating the 27 studies in the reliability sample with the content categories and the assignment criteria pertaining to them, and having him assign each study in the reliability sample to the content category which best reflected its primary content or emphasis. Analysis of the level of agreement between these assignments and those made by the present investigator, by means of Cohen's coefficient of agreement for nominal scales (Cohen, 1960), resulted in a

coefficient of agreement of .739 (see Table D, Appendix C, for the two sets of assignments). The purpose of the content categories was to provide organization for the more detailed analysis and synthesis of the individual research findings which was to take place subsequent to the large categorical analyses.

Procedure

Coding of Data. As noted in the purpose of study statement, the present investigation sought to examine findings, to be more specific, the balance of positive to negative findings, within the context of quality-of-study, source-of-derivation-of-hypotheses, measurement-scale-selected, general-design-of-study, and internal-versus-external-determination categories. Three categories were established for each variable save for measurement-scale-selected; here, there were as many categories as distinct locus of control measurements utilized across studies, and studies were coded, and assigned to categories, accordingly. With respect to the quality-of-study categories, studies receiving overall ratings of 4 or 5 were assigned to the high category; those receiving overall ratings of 3 were assigned to the medium category; and those receiving overall ratings of 2 or 1 were assigned to the low category. Three source-of-derivation-of-hypotheses categories were

identified: locus-of-control-or-attribution[1] theory, locus-of-control-or-attribution-concepts (either of which category may or may not have previous research associated with them), and previous-research-alone. Studies which made reference only to previous research in connection with their locus of control hypotheses were assigned to the previous-research-alone category; studies in which the researcher made the claim that he was testing some aspect of social learning or attribution theory, the theoretical meaning of the locus of control dimension, or such, or in which, in the absence of such a claim, the researcher discussed in some detail aspects of social learning or attribution theory, and then tied his hypothesis or hypotheses directly in to these, were assigned to the locus-of-control-or-attribution-theory category. Studies which briefly touched upon social learning or attribution theory in connection with their locus of control hypotheses, or which made use of theoretically-related but research-derived relationships in the derivation of their hypotheses, were assigned to the locus-of-control-or-attribution-concepts category. General-design-of-study categories were based upon Kerlinger's distinctions between his three major types of studies--experimental, ex post facto, and correlational, with the distinction between experimental and ex post facto being

---

[1]Social learning theory and attribution theory share the locus of control concept.

that, in the former, the researcher "has the power to assign subjects to experimental groups," and in the latter, "one cannot manipulate or assign subjects or treatments because the independent variable or variables have already occurred" (Kerlinger, 1973, p. 315). Two additional general-design-of-study categories, perhaps sub-categories would be more accurate, were developed for purposes of further analysis, a "true experimental" and a "quasi-experimental" category, based upon Campbell and Stanley's distinctions (Campbell and Stanley, 1963). It should be noted, here, though, that the sub-categories as constituted here do not correspond exactly to Campbell and Stanley's usage, the present distinction being simply one of presence or absence of random assignment in the context of a general experimental design. The purpose of this was to see whether there was any difference in balance of positive to negative findings between the "highest" type of experimental study and those accorded entry by Kerlinger's definition. Internal-versus-external-determination categories consisted of dichotomy-or-median-split, trichotomy, and extreme-groups (the latter defined by the situation in which internals and externals are drawn from the bottom and top 30% or less of the locus of control distribution). Here, studies were assigned to the category which corresponded to the method of determination employed. Only quality-of-study and general-design-of-study categories were comprehensive; in

all other instances there were studies which did not properly fit into any of the categories, or for which there was the possibility that insufficient information would be available to determine category placement. Findings of studies which did not fit a particular category were eliminated from that analysis.

With respect to the determination of whether a finding of a study was positive or not, the individual researcher's word was taken on this matter--if significance was claimed, significance was accepted, and no recalculations of original data were carried out at any point in the present study. If no indication as to significance was made, in either tables or text, the finding was not credited.

The hypothesis was the unit for the recording of findings, regardless of the number of significance tests run per hypothesis. The reason for this was to increase objectivity through avoidance of putting the present investigator in the position of having to make scores of fine distinctions as to how many relationships were really tested in a given instance, and what breakdown of the results should be adopted to reflect the most fair and accurate balance between positive and negative findings. The individual researcher's wording was the authority for how many relationships were subsumed under a single hypothesis. Separated statements of hypotheses posed no problems here; difficulties arose when a series of relationships was pre-

dicted by one statement. Here, the individual researcher was taken quite literally; in the absence of plurality wording--"hypotheses," "predictions," and such--it was assumed that all relationships in the series were subsumed under a single hypothesis. Mixed findings for single hypotheses, regardless of the proportions in relation to the individual relationships, were recorded as one finding for the positive column, one finding for the negative column.

Considerations Regarding Analysis of Findings. It was planned to calculate an overall χ^2, carrying all calculations out to four decimal points, for each of the five major category-areas, following up promising leads with partitioned χ^2s as needed. Two additional comparisons were planned to investigate possible differentials in balance of positive to negative findings for (a) true experimental versus quasi-experimental studies, and (b) for studies in which hypotheses were theory-derived versus being research-alone-derived. It was anticipated, in view of the high standing and status accorded true experimental and theory-testing research, that a greater proportion of positive findings would be found for these studies than would be the case for the other two categories, assuming, that is, that a substantial degree of construct validity existed for the locus of control construct, and that the studies examined were of reasonably high quality. Type I error probability was set at α for the collection of comparisons, to control

for the progressive increase in α that occurs with the addition of each statistical comparison (Field and Armenakis, 1974).

CHAPTER III

RESULTS

Quality-of-Study Categories

Of the 150 studies, three (2%) received overall ratings of five, 42 (28%) received overall ratings of four, 57 (38%) received overall ratings of three, 14 (9.3%) received overall ratings of two, and 34 (22.7%) received overall ratings of one (see Table E, Appendix C, for overall ratings and mean ratings assigned to each study, for assignment of studies with respect to the four categories of source-of-derivation-of-hypotheses, measurement-scale-selected, general-design-of-study, and internal-versus-external-determination, and for a tally of positive versus negative findings for each study). Examination of positive (significant) versus negative (non-significant) findings by quality-of-study categories revealed that $\underline{p} = .61$ of the findings for the low quality-of-study categories were positive, compared to $\underline{p} = .65$ of the findings for the medium quality-of-study category, and $\underline{p} = .58$ of the findings for the high quality-of-study category (findings are reported in text in this section to two decimal points, for ease of

presentation; see pertinent tables in Appendix C for complete figures). Analysis of differences in proportions across quality-of-study categories, by means of the χ^2 test statistic (α = .01, degrees of freedom = 2), yielded a non-significant overall χ^2 of .96, indicating little support for the alternative statistical hypothesis that there was a difference in proportion of positive to negative findings, beyond chance fluctuation, across quality-of-study categories. It is noteworthy, however, that the category with the lowest proportion (.58) of positive findings is the high category: that the studies rated as highest in quality produced the fewest, proportionally, positive findings. In view of the very low overall χ^2, further analysis through partitioning the contingency table was not undertaken. See Table 1, page 41, for χ^2 contingency table and mechanics of the statistical analysis.

Source-of-Derivation-of-Hypotheses Categories

Of the 150 studies, source of derivation of hypotheses could not be determined for four studies, source of derivation for nine other studies did not correspond to any of the present categories. The source-of-derivation analysis, thus, was based upon the findings of 137, rather than the full 150 studies. Examination of positive versus negative findings by source-of-derivation-of-hypotheses categories revealed that $\underline{p}$ = .54 of the findings for the

TABLE 1

Positive Versus Negative Findings
By Quality of Study Ratings

		Positive (x)	Negative (not x)	Total	Proportion
A	Low	77	49	126	.6111
B	Medium	77	42	119	.6471
C	High	59	42	101	.5842
	Total	213	133	346	

$\hat{p} = .6156$ $\hat{q} = .3844$ $\hat{p}\,\hat{q} = .2366$

$$\text{Overall } \chi^2 = \frac{\Sigma xp - \hat{p}\,T_x}{\hat{p}\,\hat{q}}$$

$$= \frac{77(.6111) + 77(.6471) + 59(.5842) - 213(.6156)}{.2366}$$

$$= \frac{47.0547 + 49.8267 + 34.4678 - 131.1228}{.2366}$$

$$= \frac{131.3492 - 131.1228}{.2366}$$

$$= \frac{.2264}{.2366} = .9569, \text{ ns } (df = 2)$$

LEGEND

$\hat{p}$ = the estimated proportion for the combined samples

$\hat{q} = 1 - \hat{p}$

previous-research-alone category were positive, compared to $\underline{p}$ = .60 of the findings for the locus-of-control-or-attribution-concepts category, and $\underline{p}$ = .69 for the locus-of-control-or-attribution-theory category. Analysis of differences in proportions across source-of-derivation categories, by means of the χ^2 statistic (α = .01, degrees of freedom = 2), yielded a non-significant overall χ^2 of 3.66, $p < .20 > .10$, indicating little support, once again, for the alternative hypothesis of a difference, beyond chance fluctuation, in proportion of positive to negative findings across source-of-derivation categories. The trend this time, however, was in the expected direction, with the highest proportion of positive findings belonging to the theory-derived hypotheses, the lowest to the previous-research-alone-derived hypotheses, but the evidence is far from compelling, failing to reach statistical trend standards (.10) even for a single statistical comparison, much less for one from a collection of comparisons.

As reported in the method section, an extra, a priori, comparison, an analysis of the difference in proportion of positive to negative findings between theory-derived and previous-research-alone-derived categories, was also planned for this area. Statistical analysis of the data, by means of χ^2 (α = .01, degrees freedom = 1), once again resulted in a non-significant finding, a χ^2 of 3.62, $p < .10 > .05$, suggesting that there is no difference in

proportion, beyond chance fluctuation, of positive to negative findings for these two categories. See Table 2, page 44, for χ^2 contingency tables and mechanics of the statistical analyses for these two comparisons.

To examine the effects of a possible uneven distribution of low quality studies obscuring a potential category effect here, the findings of studies receiving an overall one rating were eliminated from the analysis. When this was done, however, the proportions all evened out, those for theory-derived and concept-derived categories remaining in the .60s, while that for the previous-research-alone-derived category moved up into the .60s, thus reducing, rather than enhancing, the category effect. A χ^2 test was not conducted here, as it was obvious no difference in proportions across categories would be found.

Measurement-Scale-Selected Categories

In line with the findings of Prociuk and Lussier (1975), it was found that the 1966 Rotter had been used in 66% of the present studies. So dominant, in fact, was the 1966 Rotter that the next most frequently used scale in the present sample, the Intellectual Achievement Responsibility Questionnaire (IAR), was used, by contrast, in only 5.3% of the studies. Nonetheless, since the primary interest in this area of the study was the comparison of the proportion of positive to negative findings for the 1966 Rotter to

TABLE 2

Positive Versus Negative Findings by Source Of Derivation of Hypotheses

		Positive	Negative	Total	Proportion
A	Previous Research	47	40	87	.5402
B	LC Concepts	101	67	168	.6012
C	LC Theory	47	21	68	.6912
	Total	195	128	323	

$\hat{p}$ = .6037 $\hat{q}$ = .3963 $\hat{p}\,\hat{q}$ = .2392

Overall χ^2 = 3.6601 $p < .20 > .10$, ns (df = 2)

Contrast of Positive Versus Negative Findings For Hypotheses with LC Theory Versus Previous Research Alone Derivations

		Positive	Negative	Total	Proportion
A	Previous Research Alone	47	40	87	.5402
B	LC Theory	47	21	68	.6912
	Total	94	61	155	

$\hat{p}$ = .6065 $\hat{q}$ = .3935 $\hat{p}\,\hat{q}$ = .2387

Overall χ^2 = 3.6230 $p < .10 > .05$, ns (df = 1)

those of the studies that were developed in response to its perceived deficiencies, and since calculations of χ^2 with highly unequal category frequencies were presented as examples of acceptable practice in the statistical literature (Gibbons, 1976), it was decided to include in the present analysis all scales claimed to address deficiencies of the Rotter which were represented by 10 findings or more. The Mirel's factor scoring of the Rotter was also included, to see whether there was evidence of improved situational discrimination due to the changes involved in Mirel's procedure. Examination of positive versus negative findings by the measurement-scale-selected categories resulting revealed that $\underline{p} = .61$ of the findings for the Rotter category were positive, compared to $\underline{p} = .57$ of the findings for the Mirel's factor scoring category, $\underline{p} = .59$ of the findings for the IAR category, $\underline{p} = .83$ of the findings for the Nowicki-Strickland (adult form) category, and $\underline{p} = .71$ of the findings for the Levenson category. Analysis of differences in proportions across categories, by means of the χ^2 test statistic ($\alpha = .01$, degrees freedom = 4), yielded a non-significant overall χ^2 of 4.60. In this instance, however, the differences between some of the proportions looked promising, so partitioned χ^2s were calculated also, following procedures recommended by Maxwell (1961). While the overall χ^2 was not high enough to permit rejection of the null hypothesis at the .01 level through the partitioned

χ^2s, partitioned χ^2 analyses were carried out anyway, to illustrate possibilities for more detailed analysis of such data through the partitioning procedure. The primary contrast chosen was that between the Rotter and Mirel's factor categories, in the one group, and the IAR, Nowicki-Strickland, and Levenson categories in the other. The Mirel's category was combined with the Rotter category in preference to some other combination, as the 1966 Rotter scale was used in both (the difference between the two being one mainly of procedure), and the difference in proportions betwen the two was shown to be minimal anyway. This contrast resulted in a non-significant χ^2 (α = .01, degrees freedom = 1) of 2.02, suggesting that there is no difference, beyond chance fluctuation, in proportion of positive to negative findings for the Rotter and a variation of the Rotter group and a group comprised of the scales in this sample which were claimed to improve upon its deficiencies. Of course, the proportion of positive findings was higher for the Nowicki-Strickland and for the Levenson, and the combination of the IAR with these did attenuate their effect in this analysis to a certain extent, but then again, there had not been the opportunity, in this sample, at least, for those scales to accumulate the number of findings, both positive and negative, that the Rotter had. Comparison of within-group differences in proportions for the IAR, Nowicki-Strickland, and Levenson group resulted in a non-significant χ^2 of 2.79

($\alpha = .01$, degrees freedom = 2); comparison of within-group differences for the Rotter and the Mirel's factors group resulted in a non-significant χ^2 of .08 ($\alpha = .01$, degrees freedom = 1). It should be noted, here, that the outcome yielded greater differences in proportion among the "improvements" on the Rotter ($\chi^2 = 2.79$) than there was between these as a group and the two Rotter "scales" as a group ($\chi^2 = 2.02$). It should also be noted that the partitioned χ^2 values do not, in this instance, add up properly to the overall χ^2 value; this is due to rounding error derived from carrying calculations out to only four decimal places and to the very small figures involved in the latter stages of calculation of χ^2. None of the χ^2s obtained here, however, were sufficiently close to attaining significance for this to have made any difference. The elimination of the findings of studies receiving an overall one rating from the analysis would have made little difference, for, although the proportion of positive to negative findings for the Rotter category stayed about the same ($p = .60$), and those for the IAR and Levenson categories increased to $p = .65$ and $p = .80$, respectively, those for the Mirel's category rose to $p = .71$, while those for the Nowicki-Strickland category dropped to $p = .79$. See Table 3, page 48, for χ^2 contingency tables and mechanics of the statistical analyses for the four comparisons described in this section.

TABLE 3

Positive Versus Negative Findings by LC Scale Used

Scale		Positive	Negative	Total	Proportion
A	1966 Rotter	127	82	209	.6077
B	Mirel's factors	8	6	14	.5714
C	IAR	13	9	22	.5909
D	Nowicki-Strickland (Adult)	15	3	18	.8333
E	Levenson	17	7	24	.7083
	Total	180	107	287	

$\hat{p} = .6272$ $\hat{q} = .3728$ $\hat{p}\,\hat{q} = .2338$

Overall $\chi^2 = 4.5997$, ns (df = 4)

A + B	135	88	223	.6054
C + D + E	45	19	64	.7031
Total	180	107	287	

$\hat{p} = .6272$ $\hat{q} = .3728$ $\hat{p}\,\hat{q} = .2338$

Overall $\chi^2 = 2.0210$, ns (df = 1)

A	127	82	209	.6077
B	8	6	14	.5714
Total	135	88	223	

$\hat{p} = .6054$ $\hat{q} = .3946$ $\hat{p}\,\hat{q} = .2389$

Overall $\chi^2 = .0841$, ns (df = 1)

C	13	9	22	.5909
D	15	3	18	.8333
E	17	7	24	.7083
Total	45	19	64	

$\hat{p} = .7031$ $\hat{q} = .2969$ $\hat{p}\,\hat{q} = .2088$

Overall $\chi^2 = 2.7912$ ns (df = 2)

General-Design-of-Study Categories

Of the 150 studies in the present sample, 55, or 36.7%, were experimental in design, 68 (45.3%) were ex post facto in design, and 27 (18%) were correlational in design. Of the 55 experimental studies, 34 were "true experimental" in nature, while the other 21 were "quasi-experimental" in nature. Examination of the proportion of positive versus negative findings by general-design-of-study categories revealed that $\underline{p} = .67$ of the findings for the correlational-design category were positive, compared to $\underline{p} = .65$ of the findings for the ex-post-facto-design category, and $\underline{p} = .55$ of the findings for the experimental-design category. Analysis of differences in proportions across general-design-of-study categories, by means of the χ^2 test statistic ($\alpha = .01$, degrees freedom = 2), yielded a non-significant overall χ^2 of 4.12, $p < .20 > .10$. It should be noted, here, however, that the category with the highest proportion of positive findings is the correlational-design category, the category with the next highest the ex-post-facto-design category, and the category with the lowest the experimental-design category, a trend, though not of sufficient magnitude as to constitute a statistical trend, in the opposite direction from what would have been expected in this instance.

As reported in the method section, an extra, a priori comparison, an analysis of the difference in proportion of positive to negative findings between studies

classified as true experimental and those classified as quasi-experimental, was also planned for this area. Statistical analysis of the data by means of χ^2 ($\alpha = .01$, degrees freedom = 1) resulted in a significant χ^2 of 7.44, $p < .01 > .005$. The direction of the effect, however, was opposite to that anticipated; the proportion of positive findings for the true-experimental-design category, the category associated with the most highly respected research design, was significantly smaller than that for the quasi-experimental-design category (proportions of $\underline{p} = .47$ to $\underline{p} = .72$). See Table 4, page 51, for χ^2 contingency tables and mechanics of the statistical analyses for these two comparisons.

It should be noted, however, that a substantial portion of the difference above is contributed by what could be considered the, likely, spurious positive findings of low-quality quasi-experimental studies, for when the findings of studies assigned overall ratings of one are removed from the analysis, the proportion of positive findings increases to $\underline{p} = .51$ for the true experimental studies, but drops to $\underline{p} = .63$ for the quasi-experimental studies. The elimination of the findings of low quality studies from the three major categories of correlational, ex-post-facto, and experimental designs makes negligible differences in proportions, resulting in proportions of $\underline{p} = .67$, $\underline{p} = .67$, and $\underline{p} = .55$, respectively. It should be noted, however,

TABLE 4

Positive Versus Negative Findings
By General Design of Study

		Positive	Negative	Total	Proportion
A	Correlational	38	19	57	.6667
B	Ex Post Facto	105	56	161	.6522
C	Experimental	70	58	128	.5469
	Total	213	133	346	

$\hat{p}$ = .6156 $\hat{q}$ = .3844 $\hat{p}\,\hat{q}$ = .2366

Overall χ^2 = 4.1243 $p < .20 > .10$, ns (df = 2)

Positive Versus Negative Findings by True Experimental
Versus Quasi-Experimental Design of Study

		Positive	Negative	Total	Proportion
A	Quasi-Experimental	29	11	40	.7250
B	True Experimental	41	47	88	.4659
	Total	70	58	128	

$\hat{p}$ = .5469 $\hat{q}$ = .4531 $\hat{p}\,\hat{q}$ = .2478

Overall χ^2 = 7.441 $p < .01 > .005$, significant (df = 1)

that the proportion for experimental studies in general is still the lowest of the three major categories, and that for the true experimental design is still the lowest of all. See Table 4, p. 51, for χ^2 contingency tables and mechanics of the statistical analyses for these two comparisons.

Internal-versus-External-Determination Categories

Of the 150 studies, 34, or 22.7%, determined internal versus external placement by means of dichotomization, 18 (12%) by means of trichotomization, and 24 (16%) by means of extreme groups. In the cases of the other 74 studies, either an internal-external determination was not needed, or information regarding the determination procedure was not reported. Examination of the proportion of positive versus negative findings by internal-versus-external-determination categories revealed that $\underline{p} = .57$ of the findings for the dichotomy category were positive, compared to $\underline{p} = .62$ of the findings for the trichotomy category, and $\underline{p} = .63$ of the findings for the extreme-groups category. Analysis of differences in proportions of positive to negative findings across determination categories, by means of the χ^2 test statistic ($\alpha = .01$, degrees freedom = 2), resulted in a non-significant overall χ^2 of .70, suggesting that there is no difference, beyond chance fluctuation, in proportions across determination categories. The trend, here, is in the direction that might be expected, but it does not even

begin to approach statistical significance. Elimination of the findings of studies assigned an overall one rating would result in only the most negligible chances in proportions. See Table 5, page 54, for the χ^2 contingency table and mechanics of statistical analysis for this comparison.

A total of 10 χ^2 tests, all at $\alpha = .01$, were conducted in the course of the above analyses. Applying the Field and Aremenakis (1974) formula, the actual α probability of obtaining a significant result across the collection of comparisons due to chance fluctuations alone could, potentially, have run as high as .096. This was not considered an unacceptable potential α for the present purposes.

Review of Findings of Sample Studies by Content Categories

As noted in the method section, the more detailed analysis and synthesis of the individual research findings was to be facilitated through examination of areas as organized by the content categories. The basic method used here was the narrative review, though summative descriptive statistics are presented at points. Findings of studies assigned an overall one rating were not credited as valid for purposes of the following narrative review, but they were counted in proportion-of-hypotheses-supported figures. References for this section, unless otherwise noted, are to be found in Appendix A.

Locus of Control Measurements. Six studies concerned with the development of, and presentation of

TABLE 5

Positive Versus Negative Findings by Internal Versus External Determination

		Positive	Negative	Total	Proportion
A	Dichotomy	50	38	88	.5682
B	Trichotomy	29	18	47	.6170
C	Extreme Groups	33	19	52	.6346
	Total	112	75	187	

$\hat{p} = .5989$ $\hat{q} = .4011$ $\hat{p}\,\hat{q} = .2402$

Overall $\chi^2 = .6994$, ns (df = 2)

reliability and validity data concerning, new locus of control scales were assigned to this content category. A total of nine locus of control-related hypotheses (LC hypotheses) were tested by these studies; five or approximately 56% of them, were confirmed by the findings of the studies. Hypotheses which received unequivocal[2] support, both across and within studies,[3] were that an internal locus of control was positively associated with higher socio-economic status, increasing age, and minority group status (Gruen, Korte, and Baum, 1974), and that locus of control scores derived from the Nowicki-Strickland (child form) were not significantly related to measures of social desirability (Nowicki and Strickland, 1973). Other hypotheses receiving unequivocal support were that internals were likely to expend more effort when outcomes were clearly contingent upon their performance, and that beliefs regarding control over positive outcomes would mediate effort when potential outcomes were positive; beliefs regarding control over negative outcomes would mediate effort when potential outcomes were negative (Mischel, Zeiss, and Zeiss, 1974).

---

[2]What is meant by unequivocal, here, is that no mixed or negative findings were obtained for this hypothesis, or for slightly different wordings of the same hypothesis found in other studies in the same content category.

[3]This phrase, where used, makes reference to the within-category pattern of results; where deviation occurs across categories, a note will be added, in text, next to the finding affected.

Attribution of Causality. There were 11 studies concerned with the attribution of responsibility or causality for actions and outcomes; these were assigned to the present content category. A total of 24 LC hypotheses were tested by these studies; 15, or approximately 62% of them, were confirmed by the findings of the studies. Hypotheses which received unequivocal support, both across and within studies, were that (a) while both internals and externals tended to be verbally assertive in a situation where they were asked to attempt to influence someone, the internal's tendency was to be vocally assertive (assertive through intonation, "body language," and such, as opposed to words chosen) but verbally non-assertive, and the external's tendency was to be verbally assertive but vocally non-assertive (Bugenthal, Henker, and Whalen, 1976), (b) individuals with a stable estimate of their ability who were expending a usual amount of effort on a task were more likely to see disconfirmation of their expectancy regarding task outcome as being due to good or bad luck than they were to view confirmation in this manner (Feather, 1969), (c) defensive (high achievement-oriented) externals would exhibit greater decreases in expectancy following failure than congruent (low achievement-oriented) externals, and would engage in more blame projection (Hochreich, 1975), (d) the locus of the causal attribution for an outcome (attributing an outcome to ability, effort, task simplicity, or chance) would influence the effect associated

with that outcome (Riemer, 1975), and (e) while internals and externals would not differ significantly in their ascription of guilt to the defendant in a hypothetical accident situation, internals would attribute more responsibility to the defendant for the accident, feel more strongly that he should have foreseen the accident and that they would have acted differently, imagine the person to be a heavier drinker, and recommend a longer prison sentence; externals, on the other hand, would feel sorrier for him (Sosis, 1974).

Academic Achievement. There were 10 studies concerned with the relationship of locus of control to various academic achievement-related variables; these were assigned to the present category. A total of 18 LC hypotheses were tested by these studies; eight, or 44% of them, were confirmed by the findings of the studies. Hypotheses which received unequivocal support, both across and within studies, were that individuals high in resultant achievement motivation (achievement motivation minus anxiety about failure) were more likely to attribute success in an achievement situation internally (to themselves), and thus experience it as more rewarding, than individuals low in resultant achievement motivation (Weiner and Kukla, 1970), and that individuals given a chance to mediate their own reinforcers would, after a period of time, tend toward increased internality (Johnson and Croft, 1975). Other hypotheses receiving

unequivocal support were that internals tended to be more successful academically than defensive or congruent externals, that defensive externals tended to be more successful academically than congruent externals, and that female defensive externals tended to be more successful academically than male defensive externals where college students were used as subjects (Prociuk and Breen, 1975). It should be noted, here, however, that other sample studies, such as Nowicki and Duke (1974), using grade school children as subjects, and Nowicki and Strickland (1973), using grade school and secondary school subjects, failed to find unequivocal positive associations between internality and academic achievement.

Social Influence Processes. Seven studies concerned with responses to environmental manipulation and constraints were assigned to this category. A total of 14 hypotheses were tested by these studies; 10, or approximately 71% of them, were confirmed by the findings of the studies. Hypotheses which received unequivocal support, both across and within studies, were that (a) persuaders with a strong personal control orientation would attribute greater intelligence to yielders than resisters, since greater personal effectiveness was conveyed in the persuasion of an intelligent as opposed to an unintelligent individual, while persuaders with a weak personal control orientation either failed to show differential attribution or attributed

greater intelligence to the resister (Cialdini and Mirels, 1976), (b) internals would perform better in an avoidable shock condition when led to believe they had potential control over it while externals would perform better when led to believe they did not have potential control over it, and that externals would perform better than internals in a non-stress condition over which no one was led to believe they had potential control (Houston, 1972), (c) internal experimenters would induce greater changes in attitude statements on the part of their subjects than external experimenters (Phares, 1965), and (d) belief in internal control would be significantly correlated with selected indices of adjustment in a low-constraint retirement home, with an increase in degree of environmental constraint significantly modifying the relationship relative to the first setting (Wolk, 1976).

Alcoholism and Drug Abuse. Eight studies in the present sample were concerned with the relationship between locus of control and alcohol or drug use. A total of 14 hypotheses were tested in these studies; five, or approximately 37% of them, were confirmed by the findings of the studies. Hypotheses which received unequivocal support, both across and within studies, were that the mean locus of control scores of opiate addicts would be more internal than those of collegiate subjects, the differences being most salient on the Mirel's personal control factor for the

Rotter (Berzins and Ross, 1973), and that a positive relationship between locus of control and the psychasthenia and depression scales and the validity and correction scores of the MMPI would be obtained with an alcoholic population (Goss and Morosko, 1970).

Psychological Adjustment. There were 16 studies concerned with either the locus of control orientations of the psychologically maladjusted or the relationship between locus of control and factors thought to produce or be indicators of psychological adjustment or maladjustment; these were assigned to the present category. A total of 24 hypotheses were tested by these studies; 15, or approximately 62% of them, were confirmed by the findings of the studies. Hypotheses which received unequivocal support, both across and within studies, were that (a) chronic schizophrenics would be more external than normals and hospitalized controls, with hospitalized controls being more external than normals, and that these differences in locus of control orientation would be reflected in a similar variation in preferred interpersonal distance (Duke and Mullens, 1973), (b) externality was positively associated with an accumulation of stressful life events (Crandall and Lehman, 1977), (c) clients seeking counseling would be more external than a non-client group with respect to recent events, but there would be no difference in characteristic locus of control scores (Gilbert, 1976), and (d) locus of control scores were

positively correlated with depression, anxiety, and discontent, and negatively correlated with denial (Naditch, Gargan, and Micahel, 1975).

Behavior and Individual or Group Counseling. There were 10 studies concerned with the psychotherapeutic process and responses to it; these were assigned to the present category. A total of 18 hypotheses were tested by these studies; four, or approximately 22% of them, were confirmed by the findings of the studies. Hypotheses which received unequivocal support, both across and within studies, were that externals were more responsive than internals to group therapy approaches which prescribed that the leader adopt a relatively active, powerful role, while internals were more responsive than externals to less directive approaches (Abramowitz, Abramowitz, Roback, and Jackson, 1974).

Self-Control of Behavior. Three studies concerned with the relationship between locus of control and the individual's moderation of his own behavior were assigned to this category. A total of seven hypotheses were tested by these studies; two, or approximately 29% of them, were confirmed by the findings of the studies. Only one of these, however, received unequivocal support: that internals would perform significantly better than externals in increasing their EEG alpha rhythm (Johnson and Meyer, 1974).

Strategy Preferences and Learning. There were 19 studies concerned with the relationship between locus of

control and the process of learning and learning strategies; these were assigned to the present category. A total of 35 hypotheses were tested by these studies; 21, or 60% of them, were confirmed by the findings of the studies. Hypotheses which received unequivocal support, both across and within studies, were that (a) internals would show more "typical changes" in their confidence ratings following success and failure (raising them after success, lowering them after failure) than externals (Feather, 1968), and that, in a learning situation in which a child could attribute outcomes to his own skills and efforts, success would tend to enhance, failure to lead to decrements in performance for internals (Garrett and Willoughby, 1972), (b) locus of control would not be related to initial task expectancies and thus there would be no difference in original effort for internals and externals, but, when given a chance to repeat a task that was failed, internals would work harder than externals at it (Weiss and Sherman, 1973), (c) children who persisted in the face of prolonged failure were more likely than non-persisters to view outcomes, both positive and negative, as being due to their own actions, and to attribute outcomes as being due to effort more than ability (Dweck and Repucci, 1973), (d) individuals pretreated with inescapable noise would show retarded subsequent acquisition and performance, that such results would be a function of externality, and that they would be more pronounced under a skill than under

a chance set (Hiroto, 1974), (e) attractiveness of success and repulsiveness of failure would both be most pronounced in situations wherein outcomes could be attributed more to oneself than to external factors (Feather, 1967), and that, under skill conditions, internals would value self-determined rewards more than externals, and thus would avoid relying on others even though this might be to their advantage (Julian and Katz, 1968), (f) high self-esteem individuals would respond to task outcomes with a stronger belief in internal control (Epstein and Komorita, 1971) and be more receptive and responsive to both success and failure experiences than low self-esteem individuals (Ryckman, Gold, and Rodda, 1971), (g) when children were confronted with the consequences of their problem-solving behavior, failure would tend to produce higher levels of external attribution than success did (Epstein and Komorita, 1971), (h) individuals scoring low on internal locus of personal control (one of Mirel's factors) would perform better under high than low class discipline (Parent, Forward, Canter, and Mohling, 1975), and (i) internal-field independent independent individuals would perform better than external-field dependent ones on tests requiring cognitive effort and ability, and on measures of general verbal productivity (Lefcourt and Telegdi, 1971).

Interpersonal Perception and Processes. Nine studies concerned with the relationship between locus of

control and interpersonal perception and interactional processes were assigned to this category. A total of 21 hypotheses were tested by these studies; nine, or approximately 43% of them, were confirmed by the findings of the studies. Hypotheses which received unequivocal support, both across and within studies, were that (a) internals would be more attracted than externals to an individual who was a source of both positive and negative evaluation of their performance, while externals were more attracted than internals to a non-evaluative individual (Holmes and Jackson, 1975), (b) externals indicated proportionally more agreement than internals with a positive as opposed to a negative evaluator and that this relationship would be stronger in an opinion-response than an ability-response condition, as it was easier to change opinions than responses reflecting ability factors (Jones and Shrauger, 1968), (c) externals would use more coercive and fewer personal persuasive powers than internals when placed in a supervisory role (Goodstadt and Hjelle, 1973), (d) there would be no relationship between locus of control and volunteering behavior (Schwartz, 1974), and (3) internals would make no distinctions between disabled and non-disabled individuals (Navarre and Minton, 1977).

Job Competence and Satisfaction. Eight studies concerned with the relationship between locus of control and vocational aspirations, vocational counseling and train-

ing, or job competence and satisfaction were assigned to this category. A total of 27 hypotheses were tested by these studies; 17, or approximately 63% of them, were confirmed by the findings of the studies. Hypotheses which received unequivocal support, both across and within studies, were that (a) internals were more likely to see rewards as being contingent on job performance, more motivated to perform on the job, and better performers on the job than externals, and that those occupying the higher pay grades in the sample would be more internal than those occupying the lower pay grades (Broedling, 1975), (b) emotional adjustment (to include locus of control, scored in an internal direction) and organizational knowledge in a bureaucratic structure would decrease successively for managers, supervisors, and workers (Lichtman, 1970), (c) the more internal the individual, the greater would be his satisfaction with participative management, but regardless of the style of supervision, the job involvement of externals would be low (Runyon, 1973), (d) compared with their internal counterparts, women with an external orientation would aspire to work for shorter periods of their lives, be expected to have a weaker commitment to their careers, and be expected to feel more discomfort in roles that violated traditional sex-role requirements, and that externality would be associated with less support for women's liberation ideology, less positive anticipation of future careers, and less involvement in career

planning activities (Maracek and Frasch, 1977), and (e) successful internals (those who improved performance after a natural disaster) would exhibit different adjustments in locus of control than unsuccessful internals (that the locus of control of successful internals would become more internal in relation to the unsuccessful group) (Anderson, 1977).

Information Acquisition and Use. Nine studies concerned with the relationship between locus of control and the acquisition, processing, and use of situation-relevant information, and the effects of this upon subsequent response were included in this category. A total of 20 hypotheses were tested by these studies; 14, or 70% of them, were confirmed by the findings of the studies. Hypotheses which received uneuivocal support, both across and within studies, were that (a) internality is likely to be associated with a disposition toward acquiring information through one's own active and direct encounters with the enviornment, and that an increased recall period would result in a greater increment in the retention of internals than of externals (Pines, 1973), (b) internals would use the time alloted for a verbal ability test more appropriately than externals, that is, they would spend more time on the difficult, less time on the easy items of the test; no systematic pattern of time usage, on the other hand, was apparent for externals (Gozali, Cleary, Walster, and Gozali, 1973), (c) when success

feedback was conveyed by self-discovery, internals would perform better than externals, but when success feedback was conveyed through an experimenter's verbal praise, the opposite would be the case (Baron and Ganz, 1972), (d) subjects' participation in chance-related experiments would lead to an increase in feelings of internal control, while participation in skill-related experiments would lead to increases in feelings of external control: participation in experiments said to be neither chance nor skill related, however, would lead to no significant changes in locus of control scores (Eisenmann, 1972), and (e) external dental surgery patients would adjust best after viewing a general information tape regarding the dental clinic and its general procedures than a specific information tape designed to inform patients as to what to expect during the dental surgery process itself (Auerbach, Kendall, Cuttler, and Levitt, 1976).

Cross-Cultural Comparisons. Three studies concerned with the comparison of locus of control orientations across differing national groups were assigned to this category. A total of three hypotheses were tested by these studies; one, or approximately 33% of them, was confirmed by the findings of the studies. The one hypothesis in this category which received uneuivocal support was that Chinese subjects would indicate a higher preference for Rotter items emphasizing external control than would American subjects (Hsieh, Shybut, and Lotsof, 1969).

Research on the I-E Scale. Three studies concerned with the reliability, validity, and factor structure of the various locus of control scales were assigned to this category. A total of four hypotheses were tested by these studies; two, or 50% of them, were supported by the findings of the studies. Hypotheses which received unequivocal support were that locus of control as measured by the Rotter and social desirability as measured by the Edward's Social Desirability Scale were negatively correlated (Cone, 1971), and that internals, due to their expectations for being able to produce desired results through their own efforts, would be more effective in working with chronic mental patients than externals, who would feel relatively powerless to produce desired results (Hersch and Scheibe, 1967).

Social-Political Involvement. Four studies concerned with the relationship between locus of control and political participation, social-political ideology, or responses to sociopolitical events were assigned to this category. A total of four hypotheses were tested by these studies; two, or 50% of them, were supported by the findings of the studies. Hypotheses that received unequivocal support were that activist liberals would have significantly higher Powerful Others Scale scores on the Levenson than activist conservatives, as conservatives would be more in agreement with established sources of power, and that those who devoted more time and energy to feminist causes would

have stronger beliefs that powerful others exerted an influence on their lives (Levenson and Miller, 1976).

Group Differences. Seven studies concerned with the comparison of locus of control orientations of differing groups within a national group were assigned to this category. A total of 12 hypotheses were tested by these studies; five, or approximately 42% of them, were confirmed by the findings of the studies. Hypotheses which received unequivocal support, both across and within studies, were that (a) Negroes would score higher on external control than a comparable group of whites (Lefcourt and Ladwig, 1965; Zytkoskee, Strickland, and Watson, 1971), but whites would be higher on delay of gratification than Negroes (Zytkoskee, Strickland and Watson, 1971), and that Negroes would perform in a more external manner than whites on achievement tests that demanded self-evaluation (Lefcourt and Ladwig, 1965); it should be noted here, however, that another sample study, Garrett and Willoughby (1972), obtained negative results in relation to their hypothesis that lower-class black grade school children were more external than white middle-class grade school children, (b) internal Negro subjects would ask more questions in relation to a task they were to perform than external Negro subjects, but that external Negro subjects charged with the task of convincing someone of the importance of Negro contributions to American society would spend more time examining their materials than would Negro

subjects charged with the task of convincing someone of the importance of the Vice-President's role in national and foreign affairs (Williams and Stack, 1972), and (c) mental health volunteers would be more internal on the Rotter than a control group of non-volunteers (Beckman, 1972).

Personality. There were 17 articles concerned with the relationship between locus of control and various personality variables assigned to the present category. A total of 39 hypotheses were tested by these studies; 24, or approximately 62% of them, were supported by the findings of the studies. Hypotheses which received unequivocal support, both across and within studies, were that (a) there would be a negative relationship between externality and scores on the Tolor-Reznikoff Test of Insight (Tolor and Reznikoff, 1967), but that externality would be positively correlated with scores on the Authoritarian Control and Hostility Rejection factors of the Parental Attitude Research Instrument (Tolor and Jalowiec, 1968) and with acceptance of a bogus general personality interpretation (Snyder and Larson, 1972), (b) internals would show a significant reduction in response latency and increased frustration, evident in decrements in performance or increased errors, following non-rewarded trials; since the external group was predicted to be less frustrated than the internal, these effects were predicted to be of less magnitude, if present at all, for externals (Libb and Serum, 1974), (c)

externals were more anxious than internals, and the more external the individual was, the lower would be his expectation of valued-goal achievement (Strassberg, 1973); it should be noted, however, that in another sample study, Houston (1972) obtained negative results in relation to his hypotheses linking anxiety with externality, (d) externals would not only show greater preferences for tasks with built-in rationalizations than internals, but such preferences would also be associated with lower expectancies on tasks without such built-in rationalizations (Phares and Lamiell, 1974), (e) internals would report that they reacted more constructively to frustration than externals, and would show more concern with overcoming difficulties, more constructive thinking, and less preoccupation with failure than externals (Brisset and Nowicki, 1973), (f) internals would show a greater tendency to take overt remedial action to correct personal shortcomings, when provided the opportunity, than externals (Phares, Ritchie, and Davis, 1968), (g) proportionally fewer internals than externals would have experienced romantic attachments, with internals having, in addition, a less idealistic orientation toward romantic love (Dion and Dion, 1973), (h) values would be more strongly developed, and of greater clarity and relevance, in internals as opposed to externals (McKinney, 1975), (i) individuals for whom there was congruence with respect to locus of control and field independence-dependence

orientations (field independent-internal, field dependent-external) would demonstrate longer latencies and shorter durations of the effect in the autokinetic light-tracking situation, showing a more veridical interpretation of that situation, than incongruent groups; the congruent groups' correlation between their actual-self sort and their ideal-self sort on the Q-sort would also be higher, reflecting their better personality adjustment (Tobacyk, Broughton, and Vaught, 1975), and (j) internality would be positively associated with higher grade point average for males and with greater involvement in extracurricular activities for females (Nowicki and Segal, 1974).

A total of 293 hypotheses were tested by the studies in the present sample; 160, or approximately 55% of them, received unequivocal support from the findings of the studies within their respective content categories. Some of the positive findings (those obtained by studies assigned an overall one rating), were not, however, considered reliable, and were thus eliminated from further consideration, which left 124 hypotheses that received unequivocal support within categories. Three other hypotheses received unequivocal support within their own categories, but were challenged by the findings of hypotheses from studies assigned to other content categories. Of the total 293 hypotheses, 121, or approximately 41% of them, received unequivocal support both within and across all studies included in the present sample.

There were 62 hypotheses which were determined to have been derived from locus of control or attribution theory; 41, or approximately 66% of them, received unequivocal support from the findings of the studies within their respective content categories. Nine of the hypotheses that received such support were associated with studies assigned overall ratings of one, and one other was challenged by the findings of a hypothesis from a study assigned to another content category. Once these were eliminated, 31 hypotheses, or 50% of the theory-derived hypotheses examined by the present study, remained to claim unequivocal support within and across all studies in the sample.

The above figures regarding proportions of hypotheses that received unequivocal support did not provide terribly compelling evidence in favor of the locus of control construct. One might expect a rather high proportion of negative findings to be associated with a collection of weakly-derived hypotheses, but 50% support for a collection of theory-derived hypotheses, all from reasonably adequate studies, seemed low, which raised some questions regarding the construct validity of the locus of control construct.

Furthermore, results of the χ^2 comparisons were not very encouraging for the locus of control construct. The results of the χ^2 tests for proportion of positive versus negative findings by quality-of-study, source-of-derivation-of-hypotheses, measurement-scale-selected, general-design-

of-study, and internal-versus-external-determination categories indicated that there were no differences in proportion of positive to negative findings, beyond chance fluctuation, across any of the categories, for any set of categories. These results suggested that the proportion of positive to negative findings did not improve (or otherwise change) as the quality of the study, the derivation of the hypotheses, the design of the study, and the procedures for determining whether an individual's locus of control orientation was most properly considered an internal or an external one did. Also suggested by these findings was the fact that there was little evidence, in the present sample, that the LC measurement scales claimed to improve upon the Rotter instrument produced any higher proportion of positive to negative findings than did the Rotter itself. Even the supplemental, direct comparison of the proportion of positive to negative findings between theory-derived and previous-research-alone-derived categories failed to yield a significant finding. The one significant finding obtained, i.e., the comparison of proportion of positive to negative findings between studies classified as true experimental or quasi-experimental, yielded a result opposite to that anticipated, indicating a significantly <u>smaller</u>, rather than larger, proportion of positive findings for the studies classified as true experimental in design.

CHAPTER IV

SUMMARY, CONCLUSIONS, DISCUSSION, AND IMPLICATIONS

Summary

The research literature in psychology and education had expanded at an astounding rate in recent years (Glass, 1976; Glass, 1977; Kulik, Kulik, and Cohen, 1979), but the traditional narrative review was considered incapable of organizing, analyzing, and synthesizing the vast bodies of information that had accumulated (Glass, 1976). An improved, more systematic and objective method of review, the meta-analysis of research, was described by Glass (1976), but it had been used only with collections of studies that were homogeneous with respect to major variables measured, and it assigned equal weight to the findings of studies which varied widely with respect to important sampling, design, and analysis considerations. The present study investigated possibilities for a meta-analysis of the findings of a collection of studies which were heterogenous with respect to major variables measured, and took into account, at the same time, variations in sampling, data-collection, design, and analysis across studies. The locus of control research literature

was selected for the analysis, as research had been expanding rapidly in this area, but objective, systematic analyses had not been undertaken. Objectives specific to an analysis of the locus of control research area included (a) an investigation of the construct validity of the locus of control construct through an examination of findings within the context of sources-of-derivation-of-hypotheses categories, (b) an investigation of possible differentials related to locus of control measurement procedures through an examination of findings within the context of measurement-scale-selected categories, (c) an investigation of possible differentials related to design of study through an examination of findings within the context of general-design-of-study categories, and (d) an investigation of possible differentials related to criteria for locus of control group-assignment through an examination of findings within the context of internal-versus-external-determination categories.

Since a systematic review of more than 2200 locus of control studies was not considered feasible, analysis was confined to a sample of the population of locus of control studies. Since the journals of the American Psychological Association provided a broader coverage of the various areas and specialties of psychology than those of any other publisher, and since the APA journals published rather sizable proportions of the locus of control studies published in journals, it was decided that all locus of

control studies appearing in APA journals from 1957 through 1977 which obtained a measure of locus of control orientation, tested hypotheses concerned with a locus of control issue, and were reported in articles greater than one page in length would be included in the present investigation. Studies were located by means of a search of Psychological Abstracts and of several pertinent reviews and bibliographies. No studies published before 1963 qualified for the final sample. The final sample consisted of 150 studies.

The levels of quality of the studies in relation to various sampling, data-collection, design, and analysis considerations were assessed by means of the Quality of Study Rating Form, developed especially for the present study. The Quality of Study Rating Form, based upon Wandt's Form for the Evaluation of an Article (Wandt, 1967), and a review of several evaluation criteria sources, was originally intended to be used without specific scoring criteria, but poor reliability in a pilot assessment led to the development of specific scoring criteria. The reliability of the revised form was tested by presenting three psychologists familiarized with the assessment criteria a randomly selected sample of articles to be rated. The extent of agreement between these ratings and the ratings of the same studies by the present investigator was assessed by means of the r_{11} coefficient (Lindquist, 1953). A coefficient of .73 was obtained.

The objectives of the study were addressed through an examination of findings, to be more specific, the balance of positive to negative findings, within the context of quality-of-study, source-of-derivation-of-hypotheses, measurement-scale-selected, general-design-of-study, and internal-versus-external-determination. Criteria for assignment of studies to categories were discussed in detail. Differences in proportion of positive to negative findings across categories were assessed by means of the χ^2 test statistic. An overall χ^2 was calculated for each of the five major category areas, and promising leads were followed up with partitioned χ^2s, as needed. Two additional comparisons were made to investigate possible differentials in balance of positive to negative findings for (a) true experimental versus quasi-experimental studies, and (b) for studies in which hypotheses were theory-derived versus previous-research-alone-derived. It was anticipated, in view of the high standing and status accorded true experimental and theory-testing research, that a greater proportion of positive findings would be found for these studies than would be the case for the other two categories.

A total of 10 χ^2 tests were conducted. Type I error probability was set at $\alpha = .10$ for the collection of comparisons, to control for the progressive increase in α that had been shown to occur with the addition of each statistical comparison (Field and Armenakis, 1974; Petrinovich and Hardyck,

1969). Of the 10 comparisons, only one, the comparison of proportion of positive to negative findings for true experimental versus quasi-experimental studies, attained significance at the $p < .01$ level. Contrary to expectation, however, the direction of results was opposite to that anticipated, with the true experimental studies having the smaller proportion of positive findings.

A narrative and descriptive-statistical review of findings receiving unequivocal support across the sample of studies was presented in addition to the examination of differences in proportion of positive to negative findings across the five category areas. Studies were assigned to content categories, and analyses of findings were then conducted within the context of the categories to which the studies were assigned. The content categories utilized were those developed by Prociuk and Lussier (1975). As Prociuk and Lussier had not developed explicit assignment criteria for their categories (Prociuk, 1979), it was necessary for the present investigator to develop his own. This was accomplished by means of an induction of descriptors obtained through a review of the abstracts of all articles included in the study. Studies were assigned to content categories based, however, upon examination of the entire text of the article. The reliability of the assignment criteria was tested by familiarizing one of the psychologists rating the studies in the reliability sample with the content categories and the assignment criteria pertaining to them, and

having him assign each study in the reliability sample to the content category which best reflected its primary content. Analysis of the level of agreement between these assignments and those made by the present investigator resulted in a coefficient of agreement (Cohen, 1960) of .702. The review presented percentages of hypotheses confirmed by the findings of the studies and hypotheses which received unequivocal support both within and across content categories, for each of the 17 Prociuk and Lussier content categories.

Conclusions

The findings from this investigation seemed to warrant the following conclusions:

1. The general concepts and procedures of the meta-analysis of research method may be productively applied to the review of collections of studies heterogeneous with respect to variables measured, as well as to collections of studies homogeneous with respect to variables measured.

2. The reliability of study or article rating-scales may be enhanced by the development and utilization of specific scoring criteria.

3. The construct validity of the locus of control construct received little support from the quantitative analyses conducted in the present study.

Discussion

It was anticipated that a greater proportion of

positive findings would be associated with true experimental and theory-testing research than with quasi-experimental research and research which tested hypotheses derived from previous studies alone. This did not prove to be the case, however. In addition, no differences in proportion of positive to negative findings were found across quality-of-study, source-of-derivation-of-hypotheses, measurement-scale-selected, general-design-of-study, and internal-versus-external-determination categories. The one statistical test with which a significant result was associated revealed an outcome opposite to that anticipated, that there was a significantly smaller proportion of positive findings for true experimental research than for quasi-experimental research.

While the quantitative analyses discussed above, to include percentage-of-hypotheses-receiving-unequivocal-support figures, provided little support for the construct validity of the locus of control construct, these data did not reveal the entire situation, as many of the hypotheses in the theory-derived category were assigned there due to the individual researcher's claim or to a presented tie-in with aspects of Rotter's theory (or attribution theory), but merely investigated the various and sundry hypothesized characteristics of internals and externals, loosely based upon social learning theory or various relationships supposedly derived from it. Few hypotheses investigated actual processes described in social learning theory. An examina-

tion of those that did may help to shed further light upon this matter.

Social learning theory stated that reinforcement alone did not adequately explain behavior, that an individual must have had an expectancy that the behavior in question would result in the attainment of the goal he was working toward before it would be emitted, and further, that a given expectancy was a function of an associated specific expectancy (which was based upon previous experience that a valued reinforcement would or would not occur in the situation), and of various generalized expectancies, of which locus of control was one, the status of which were likewise derived from previous experiences. This process was considered central to social learning theory, and to the relationship between it and the locus of control construct. Several studies, presenting both positive and negative findings, addressed this process.

Weiner, Heckhausen, Meyer, and Cook (1972), proceeding from the perspective of attribution theory, argued that earlier positive locus of control findings regarding shifts in expectancy had been contaminated by the stability of the attributions measured, and demonstrated, in their own study, that it was the stability of the attribution, whether outcomes were attributed to the fixed factors of ability or task difficulty as opposed to the variable factors of effort or luck, rather than perceived locus of control of reinforce-

ment, that mediated expectancy. The Weiner, et al., study was assigned an overall three rating, and it was difficult to find any serious fault with it; it was, in addition, bolstered by a finding from Miller and Seligman (1973) that internals were no more likely to change expectancies following reinforcement than externals (it was argued that since externals did not perceive themselves as being able to reliably produce reinforcement for themselves, but rather saw reinforcement as being response-independent, they would be less likely to change expectancies than internals, who did see themselves as capable of exercising control over such events), and, the Miller and Seligman study, too, was a quite adequate study. Arrayed against these, on the positive side of the ledger, this time, were findings of studies conducted by Epstein and Komorita (1971), Gilbert (1976), and Hiroto (1974), all suggesting that subjection to situations in which the individual was deprived of, or rendered unable to obtain, reinforcement was productive of an external locus of control orientation, just as social learning theory would predict they would be (successful past experience with a given behavior in a given situation was considered to lead to the expectancy that it would probably work again in that same, and related, situations in the future, and thus, potentially, bring the reinforcement under the individual's control; unsuccessful experience would reduce the expectancy that the behavior would lead to the

attainment of the goal, and that the reinforcement was potentially under the individual's control). Further support was provided by Strassberg (1973), who found that the more external an individual was, the lower was his expectation of valued-goal achievement. Finally, Riemer (1975) obtained negative results in relation to his hypothesis that (the same, essentially, as the Weiner, et al., hypothesis) the stability of the causal attribution would determine the expectancy response. All the positive studies here were, in addition, reasonably adequate studies, too. In sum, then, the findings of the present sutdy would have to be regarded as mixed; little support for the construct validity of the locus of control construct was found in the results of the quantitative analysis, but the balance of findings from the studies that most closely addressed predictions central to social learning theory was favorable.

Implications

One of the major purposes of the present study was to investigate possibilities for the meta-analysis of collections of studies heterogeneous with respect to variables measured; while the present analytic and measurement procedures were rather crude, as is often the case with initial ventures into a new area of application, it was felt that the present study illustrated the possibility of, and some possibilities for, the meta-analysis of such collections of

studies. In addition, it was felt that the study provided some indication of possibilities for the improvement of the reliability of article and study rating-scales through development and utilization of item-by-item scoring criteria (it is recognized that the criteria do not provide comprehensive coverage of all potentially important scoring points, but, though moderate, the scale's reliability is a substantial improvement over that of most scales in this area), and for the importance of conducting a detailed examination of a study before accepting its findings.

It is recommended that future research refine both such item-by-item scoring criteria approaches to article or study rating-scales, and the meta-analysis procedures for the review of collections of studies heterogeneous with respect to variables measured that were presented here. It is also recommended that more theory-testing research and fewer investigations of the possible characteristics of "internals" and "externals" be undertaken in the locus of control research area, and that the present study be replicated, either by an examination of other large-scale samples of the locus of control literature utilizing the present categories, instruments, criteria, and procedures, or by an examination of the present sample utilizing other categories, instruments, criteria, and procedures, to see if the current findings are indeed representative.

REFERENCES

Abrahamson, D., Schulderman, S., and Schulderman, E. Replication of dimensions of locus of control. Journal of Consulting and Clinical Psychology, 1973, 41, 320.

American Educational Research Association. (Form used by AERA reviewers of manuscripts, date unknown.)

American Psychological Association. 1979 APA membership register. Washington, D.C.: Author, 1979.

American Psychological Association. 1978 directory of the APA. Washington, D.C.: Author, 1978.

American Psychological Association. Manuscript rating form. Washington, D.C.: Author, date unknown.

Anastasi, A. Psychological testing. London: MacMillan, 1968.

Campbell, D. and Stanley, J. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1963.

Cohen, J. A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 1960, 20, 37-46.

Collins, B. Four separate components of the Rotter I-E Scale: Belief in a difficult world, a just world, a predictable world, and a politically responsive world. Journal of Personality and Social Psychology, 1974, 29, 381-391.

Cook, T. and Gruder, C. Metaevaluation research. Evaluation Quarterly, 1978, 2, 5-51.

Field, H. and Armenakis, A. On use of multiple tests of significance in psychological research. Psychological Reports, 1974, 35, 427-431.

Galfo, A. Interpreting educational research. Dubuque, Iowa: William C. Brown, 1975.

Gibbons, J. Nonparametric methods for quantitative analysis. New York: Holt, Rinehart, and Winston, 1976.

Glass, B. and Norwood, S. How scientists actually learn of work important to them. Proceedings of the International Conference on Scientific Information (Area 1), 1958, 185-187.

Glass, G. Primary, secondary, and meta-analysis of research. Educational Researcher, 1976, 5, 3-8.

Glass, G. Integrating findings: The meta-analysis of research. Review of Research in Education, 1977, 5, 351-379.

Guilford, J. Fundamental statistics in psychology and education. New York: McGraw-Hill, 1965.

Gurin, P., Gurin, G., Lao, R., and Beattie, M. Internal-external control in the motivational dynamics of Negro youth. Journal of Social Issues, 1969, 25, 29-53.

Hopkins, K. and Chadbourn, R. A schema for proper utilization of multiple comparisons in research and a case study. American Educational Research Journal, 1967, 4, 407-412.

Isaac, S. and Michael, W. Handbook in research and evaluation. San Diego: Robert K. Knapp, 1971.

James, W. Internal versus external control of reinforcement as a basic variable in learning theory. Unpublished doctoral dissertation, Ohio State University, 1957.

James, W., Woodruff, A., and Werner, W. Effect of internal and external control upon changes in smoking behavior. Journal of Consulting Psychology, 1965, 29, 184-186.

Joe, V. Review of the internal-external control construct as a personality variable. Psychological Reports, 1971, 28, 619-640. (Monograph Supplement 3-V28).

Kerlinger, F. Foundations of behavioral research (2nd ed.). New York: Holt, Rinehart, and Winston, 1973.

Kirk, R. Experimental design: Procedures for the behavioral sciences. Belmont, California: Wadsworth, 1968.

Kulik, J., Kulik, C., and Cohen, P. A meta-analysis of outcome studies of Keller's personalized system of instruction. American Psychologist, 1979, 34, 307-318.

Lefcourt, H. Internal versus external control of reinforcement. Psychological Bulletin, 1966, 65, 206-220.

Lefcourt, H. Recent developments in the study of locus of control. In B. Maher (Ed.), Progress in experimental personality research. Vol. 6. New York: Academic Press, 1972, 1-39.

Levenson, H. and Miller, J. Multidimensional locus of control in socio-political activists of conservative and liberal ideologies. Journal of Personality and Social Psychology, 1976, 33, 199-208.

Lichtenstein, E. and Keutzer, C. Further normative and correlational data on the Internal-External (I-E) Control of Reinforcement Scale. Psychological Reports, 1967, 21, 1014-1016.

Lindquist, E. Design and analysis of experiments in psychology and education. Boston: Houghton Mifflin, 1953.

MacDonald, A. Internal-external locus of control: A partial bibliography (II). JSAS Catalog of Selected Documents in Psychology, 1972, 2, 68.

MacDonald, A. and Davis, A. Internal-external locus of control: A partial bibliography (III). JSAS Catalog of Selected Documents in Psychology, 1974, 4, 44.

Mansfield, R. and Busse, T. Meta-analysis of research: A rejoinder to Glass. Educational Researcher, 1977, 6, 3.

Maxwell, A. Analysing qualitative data. New York: John Wiley, 1961.

Mirels, H. Dimensions of internal versus external control. Journal of Consulting and Clinical Psychology, 1970, 34, 226-228.

Petrinovich, L. and Hardyck, C. Error rates for multiple comparison methods: Some evidence concerning the frequency of erroneous conclusions. Psychological Bulletin, 1969, 71, 43-54.

Phares, E. Changes in expectancy in skill and chance situations. Unpublished doctoral dissertation, Ohio State University, 1955.

Phares, E. Locus of control: A personality determinant of behavior. Morristown, New Jersey: General Learning Press, 1973.

Phares, E. Locus of control in personality. Morristown, New Jersey: General Learning Press, 1975.

Prociuk, T. Personal communication, May, 1979.

Prociuk, T. and Lussier, R. Internal-external locus of control: An analysis and bibliography of two years of research (1973-1974). Psychological Reports, 1975, 37, 1323-1337. (Monograph Supplement 1-V37).

Reid, D. and Ware, E. Multidimensionality of internal versus external control: Addition of a third dimension and nondistinction of self versus others. Unpublished manuscript, University of Waterloo, 1973.

Rohsenow, D. and O'Leary, M. Locus of control research on alcoholic populations: A review. I. Development, scales, and treatment. The International Journal of the Addictions, 1978, 13, 55-78.

Rotter, J. Social learning and clinical psychology. Englewood Cliffs, New Jersey: Prentice-Hall, 1954.

Rotter, J. Generalized expectancies for internal versus external control of reinforcement. Psychological Monographs, 1966, 80, (1, Whole Number 609).

Rotter, J. Some problems and misconceptions related to the construct of internal versus external control of reinforcement. Journal of Consulting and Clinical Psychology, 1975, 43, 56-67.

Schneider, J. and Parsons, O. Categories on the Locus of Control Scale and cross-cultural comparisons in Denmark and the United States. Journal of Cross-Cultural Psychology, 1970, 1, 131-138.

Smith, M. and Glass G. Meta-analysis of psychotherapy outcome studies. American Psychologist, 1977, 32, 752-760.

Straits, B. and Sechrest, L. Further support of some findings about the characteristics of smokers and non-smokers. Journal of Consulting Psychology, 1963, 27, 282.

Thornhill, M., Thornhill, G., and Youngman, M. A computerized and categorized bibliography on locus of control. Psychological Reports, 1975, 36, 505-506.

Throop, W. and MacDonald, A. Internal-external locus of control: A bibliography. Psychological Reports, 1971, 28, 175-190. (Monograph Supplement 1-V28).

Van Dalen, D. Understanding educational research: An introduction (3rd ed.). New York: McGraw-Hill, 1973.

Walker, R., Nast, E., Chaney, E., and O'Leary, M. Changes in drinking-related locus of control as a function of length of alcoholism treatment. Psychological Reports, 1979, 44, 287-293.

Wallston, B., Wallston, K., Kaplan, G., and Maides, S. Development and validation of the Health Locus of Control (HLC) Scale. Journal of Consulting and Clinical Psychology, 1976, 44, 580-585.

Wandt, E. Report of the (AERA) committee on evaluation of research (E. Wandt, chairman). Unpublished manuscript, 1967. (Available from E. Wandt, 670 Hamlet, Monterey Park, California, 91754).

White, K. The relationship between socioeconomic status and academic achievement. Unpublished doctoral dissertation, University of Colorado at Boulder, 1976.

Wiersma, W. Research methods in education: An introduction. New York: J. B. Lippincott, 1969.

Wyatt, W. and Bridges, C. Statistics for the behavioral sciences. Boston: D. C. Heath, 1967.

APPENDIX A

LIST OF SAMPLE ARTICLES, WITH NUMERICAL-ALPHABETICAL REFERENCE CODINGS FOR SAMPLE STUDIES

Study Codings	Articles
1	Abramowitz, C., Abramowitz, S., Roback, H., and Jackson, C. Differential effectiveness of directive and nondirective group therapies as a function of client internal-external control. Journal of Consulting and Clinical Psychology, 1974, 42, 849-853.
2	Allen, G., Giat, L., and Cherney, R. Locus of control, test anxiety, and student performance in a personalized instruction course. Journal of Educational Psychology, 1974, 66, 968-973.
3	Anderson, C. Locus of control, coping behaviors, and performance in a stress setting: A longitudinal study. Journal of Applied Psychology, 1977, 62, 446-451.
4	Auerbach, S., Kendall, P., Cuttler, H., and Levitt, N. Anxiety, locus of control, type of preparatory information, and adjustment to dental surgery. Journal of Consulting and Clinical Psychology, 1976, 44, 809-818.
5	Baron, R. and Ganz, R. Effects of locus of control and type of feedback on the task performance of lower-class black children. Journal of Personality and Social Psychology, 1972, 21, 124-130.
6	Beckman, L. Locus of control and attitudes toward mental illness among mental health volunteers. Journal of Consulting and Clinical Psychology, 1972, 38, 84-89.
7	Berman, A. and Hays, J. Relation between death anxiety, belief in afterlife, and locus of control. Journal of Consulting and Clinical Psychology, 1973, 41, 318.
8	Berzins, J. and Ross, W. Locus of control among opiate addicts. Journal of Consulting and Clinical Psychology, 1973, 40, 84-91.
9	Bigoness, W. Effects of locus of control and style of third party intervention upon bargaining behavior. Journal of Applied Psychology, 1976, 61, 305-312.
10	Brandt, J. Internal versus external locus of control and performance in controlled and motivated reading-rate improvement instruction. Journal of Counseling Psychology, 1975, 22, 377-383.
11	Brandt, L., Hayden, M., and Brophy, J. Teacher's attitudes and ascription of causation. Journal of Educational Psychology, 1975, 67, 677-682.
12	Brissett, M. and Nowicki, S. Internal versus external control of reinforcement and reaction to frustration. Journal of Personality and Social Psychology, 1973, 25, 35-44.

13 Broedling, L. Relationship of locus of control to work motivation and performance in an expectancy model. Journal of Applied Psychology, 1975, 60, 65-70.

14 Bryant, B. and Trockel, J. Personal history for psychological stress related to locus of control orientation among college women. Journal of Consulting and Clinical Psychology, 1976, 44, 266-271.

15 Bugenthal, D, Henker, B. and Whalen, C. Attributional antecedents of verbal and vocal assertiveness. Journal of Personality and Social Psychology, 1976, 34, 405-411.

16 Caster, D. and Parsons, O. Relationship of depression, sociopathy, and locus of control to treatment outcome in alcoholics. Journal of Consulting and Clinical Psychology, 1977, 45, 751-756.

17 Cherulnik, P. and Citrin, M. Individual difference in psychological reactness: The interaction between locus of control and mode of elimination of freedom. Journal of Personality and Social Psychology, 1974, 29, 398-404.

18 Cialdini, R. and Mirels, H. Sense of personal control and attributions about yielding and resisting persuasion targets. Journal of Personality and Social Psychology, 1976, 33, 395-402.

19 Cochran, D., Hoffman, D., Strand, K., and Warren, P. Effects of client/computer interaction on career decision-making processes. Journal of Counseling Psychology, 1977, 24, 308-312.

20a,b Cohen, D. and Wolfe, G. Dream recall and repression: Evidence for an alternative hypothesis. Journal of Consulting and Clinical Psychology, 1973, 41, 349-355.

21 Cone, J. Locus of control and social desirability. Journal of Consulting and Clinical Psychology, 1971, 36, 449

22 Crandall, J. and Lehman, R. Relationship of stressful life events to social interest, locus of control, and psychological adjustment. Journal of Consulting and Clinical Psychology, 1977, 45, 1208.

23 Diamond, M. and Shapiro, J. Changes in locus of control as a function of encounter group experiences. Journal of Abnormal Psychology, 1973, 82, 514-518.

24 Dion, K. L. and Dion, K. K. Correlates of romantic love. Journal of Consulting and Clinical Psychology, 1973, 41, 51-56.

25 Dua, P. Comparisonof the effects of behaviorally oriented action and psychotherapy reeducation on introversion-extraversion, emotionality, and internal-external control. Journal of Counseling Psychology, 1970, 17, 567-572.

26 DuCette, J. and Wolk, S. Cognitive and motivational correlates of generalized expectancies for control. Journal of Personality and Social Psychology, 1973, 26, 420-426.

27 Duke, M. and Mullens, C. Preferred interpersonal distance as a function of locus of control orientation in chronic schizophrenics, nonschizophrenic patients, and normals. Journal of Consulting and Clinical Psychology, 1973, 41, 230-234.

28a,b Duncan, C. A reputation test of personality integration. Journal of Personality and Social Psychology, 1966, 3, 516-524.

29 Dweck, C. and Reppucci, N. Learned helplessness and reinforcement responsibility in children. Journal of Personality and Social Psychology, 1973, 25, 109-116.

30 Eisenman, R. Experience in experiments and change in internal-external control scores. Journal of Consulting and Clinical Psychology, 1972, 39, 434-435.

31 El-Gazzar, M., Saleh, S., and Conrath, D. Situational and individual difference variables in chance-determined activities. Journal of Personality and Social Psychology, 1976, 34, 951-959.

32 Epstein, R. and Komorita, S. Self-esteem, success-failure, and locus of control in Negro children. Developmental Psychology, 1971, 4, 2-8.

33 Feather, N. Valence of outcome and expectation of success in relation to task difficulty and perceived locus of control. Journal of Personality and Social Psychology, 1967, 7, 372-386.

34 Feather, N. Change in confidence following success or failure as a predictor of subsequent performance. Journal of Personality and Social Psychology, 1968, 9, 38-46.

35 Feather, N. Attribution of responsibility and valence of success and failure in relation to initial confidence and task performance. Journal of Personality and Social Psychology, 1969, 13, 129-144.

36 Fontana, A., Klein, E., Lewis, E., and Levine, L. Presentation of self in mental illness. Journal of Consulting and Clinical Psychology, 1968, 32, 110-119.

37 Fry, P. Interaction between locus of control, level of inquiry, and subject control in the helping process: A laboratory analogue study. Journal of Counseling Psychology, 1975, 22, 280-287.

38 Garrett, A. and Willoughby, R. Personal orientation and reactions to success and failure in urban black children. Developmental Psychology, 1972, 7, 92.

39a,b Gilbert, L. Situational factors and the relationship between locus of control and psychological adjustment. Journal of Counseling Psychology, 1976, 23, 302-309.

40 Goodstadt, B. and Hjelle, L. Power to the powerless: Locus of control and the use of power. Journal of Personality and Social Psychology, 1973, 27, 190-196.

41 Goss, A. and Morosko, T. Relation between a dimension of internal-external control and the MMPI with an alcoholic population. Journal of Consulting and Clicical Psychology, 1970, 34, 189-192.

42 Gozali, H., Cleary, A., Walster, W., and Gozali, J. Relationship between the internal-external control construct and achievement. Journal of Educational Psychology, 1973, 64, 9-14.

43 Gruen, G., Korte, J., and Baum, J. Group measure of locus of control. Developmental Psychology, 1974, 10,683-686.

44 Hamsher, J., Geller, J., and Rotter, J. Interpersonal trust and the Warren Commission Report. Journal of Personality and Social Psychology, 1968, 9, 210-215.

45 Hersch, P. and Scheibe, K. Reliability and validity of internal-external control as a personality dimension. Journal of Consulting Psychology, 1967, 31, 609-613.

46 Higgins, R. and Marlatt, G. Fear of interpersonal evaluation as a determinant of alcohol consumption in male social drinkers. Journal of Abnormal Psychology, 1975, 84, 644-651.

47 Hiroto, D. Locus of control and learned helplessness. Journal of Experimental Psychology, 1974, 102. 187-193.

48 Hochreich, D. Defensive externality and blame projection following failure. Journal of Personality and Social Psychology, 1975, 32, 540-546.

49 Holmes, D. and Jackson, T. Influence of locus of control on interpersonal attraction and affective reactions in situations involving reward and punishment. Journal of Personality and Social Psychology, 1975, 31, 132-136.

50 Houston, B. Control over stress, locus of control, and response to stress. Journal of Personality and Social Psychology, 1972, 21, 249-255.

51 Howie, A. Effects of brief exposure to symbolic model behavior on the information-processing strategies of internally and externally-oriented children. Developmental Psychology, 1975, 11, 325-333.

52 Hrycenko, I. and Minton, H. Internal-external control, power position, and satisfaction in task-oriented groups. Journal of Personality and Social Psychology, 1974, 30, 871-878.

53 Hsieh, T., Shybut, J., and Lotsof, E. Internal versus external control and ethnic group membership: A cross-cultural comparison. Journal of Consulting and Clinical Psychology, 1969, 33, 122-124.

54 James, W., Woodruff, A., and Werner, W. Effect of internal and external control upon changes in smoking behavior. Journal of Consulting Psychology, 1965, 29, 184-186.

55 Jessor, R., Young, H., Young, E., and Tesi, G. Perceived opportunity, alienation, and drinking behavior among Italian and American youth. Journal of Personality and Social Psychology, 1970, 15, 215-222.

56 Johnson, C. and Gormly, J. Academic cheating: The contribution of sex, personality, and situational variables. Developmental Psychology, 1972, 6, 320-325.

57 Johnson, J., Leventhal, H., and Dabbs, J. Contribution of emotional and instrumental response processes in adaptation to surgery. Journal of Personality and Social Psychology, 1971, 20, 55-64.

58 Johnson, R. and Meyer, R. The locus of control construct in EEG alpha rhythm feedback. Journal of Consulting and Clinical Psychology, 1974, 42, 913.

59 Johnson, W. and Croft, R. Locus of control and participation in a personalized system of instruction course. Journal of Educational Psychology, 1975, 67, 416-421.

60 Jones, S. and Shrauger, J. Locus of control and interpersonal evaluations. Journal of Consulting and Clinical Psychology, 1968, 32, 664-668.

61a,b Julian, J. and Katz, S. Internal versus control and the value of reinforcement. Journal of Personality and Social Psychology, 1968, 8, 89-94.

62 Karabenick, S. Valence of success and failure as a function of achievement motives and locus of control. Journal of Personality and Social Psychology, 1972, 21, 101-110.

63a,b Kilmann, P., Albert, B., and Sotile, W. Relationship between locus of control, structure of therapy, and outcome. Journal of Consulting and Clinical Psychology, 1975, 43, 588.

64 Kilmann, P. and Howell, R. Effects of structure of marathon group therapy and locus of control on therapeutic outcome. Journal of Consulting and Clinical Psychology, 1974, 42, 912.

65a,b, and c Korman, A. Environmental ambiguity and locus of control as interactive influences on satisfaction. Journal of Applied Psychology, 1971, 55, 339-342.

66 Lao, R. Internal-external control and competent and innovative behavior among Negro college students. Journal of Personality and Social Psychology, 1970, 14, 263-270.

67 Lefcourt, H. Risk taking in Negro and white adults. Journal of Personality and Social Psychology, 1965, 2, 765-770.

68 Lefcourt, H. Effects of cue explication upon persons maintaining external control expectancies. Journal of Personality and Social Psychology, 1967, 5, 372-378.

69 Lefcourt, H. and Ladwig, G. The American Negro: A problem in expectancies. Journal of Personality and Social Psychology, 1965, 1, 377-380.

70 Lefcourt, H. and Telegdi, M. Perceived locus of control and field dependence as predictors of cognitive activity. Journal of Consulting and Clinical Psychology, 1971, 37, 53-56.

71 Levenson, H. Perceived parental antecedents of internal powerful others, and chance locus of control orientations. Developmental Psychology, 1973, 9, 260-265.

72a,b Levenson, H. and Miller, J. Multidimensional locus of control in sociopolitical activists of conservative and liberal ideologies. Journal of Personality and Social Psychology, 1976, 33, 199-208.

73 Libb, J. and Serum, C. Reactions to frustrative nonreward as a function of perceived locus of control of reinforcement. Journal of Experimental Psychology, 1974, 102, 494-497.

74 Lichtman, C. Some intrapersonal correlates of organizational rank. Journal of Applied Psychology, 1970, 54, 77-80.

75 Lied, T. and Pritchard, R. Relationships between personality variables and components of the expectancy-valence model. Journal of Applied Psychology, 1976, 61, 463-467.

76 Lipp, L., Kolstoe, R., James, W., and Randall, H. Denial of disability and internal control of reinforcement: A study using a perceptual defense paradigm. Journal of Consulting and Clinical Psychology, 1968, 32, 72-75.

77 Lottman, T. and DeWolfe, A. Internal and external control in reactive and process schizophrenia. Journal of Consulting and Clinical Psychology, 1972, 39, 344.

78 MacDonald, A. and Hall, J. Perception of disability by the non-disabled. Journal of Consulting and Clinical Psychology, 1969, 33, 654-660.

79 MacDonald, A. and Hall, J. Internal-external locus of control and perception of disability. Journal of Consulting and Clinical Psychology, 1971, 36, 338-343.

80 Marecek, J. and Frasch, C. Locus of control and college women's role expectations. Journal of Counseling Psychology, 1977, 24, 132-136.

81 McArthur, L. Luck is alive and well in New Haven: A serendipitous finding on perceived control of reinforcement after the draft lottery. Journal of Personality and Social Psychology, 1970, 16, 316-318.

82a,b McKinney, J. The development of values: A perceptual interpretation. Journal of Personality and Social Psychology, 1975, 31, 801-807.

83 McMahan, I. Relationship between causal attributions and expectancy of success. Journal of Personality and Social Psychology, 1973, 28, 108-114.

84a,b Miller, W. and Arkowitz, H. Anxiety and perceived causation in social success and failure experiences: Disconfirmation of an attribution hypothesis in two experiments. Journal of Abnormal Psychology, 1977, 86, 665-668.

85 Miller, W. and Seligman, M. Depression and the perception of reinforcement. Journal of Abnormal Psychology, 1973, 82, 62-73.

86 Mischel, W., Zeiss, R., and Mischel, A. Internal-external control and persistence: Validation and implications of the Stanford Preschool Internal-External Scale. Journal of Personality and Social Psychology, 1974, 29, 265-278.

87 Naditch, M. and Fenwick, S. LSD flashbacks and ego functioning. Journal of Abnormal Psychology, 1977, 86, 352-359.

88 Naditch, M., Gargan, M., and Michael, L. Denial, anxiety, locus of control, and discrepancy between aspirations and achievements as components of depression. Journal of Abnormal Psychology, 1975, 84, 1-9.

89 Navarre, K. and Minton, H. Internal-external control and attitude toward disability. Journal of Consulting and Clinical Psychology, 1977, 45, 961-962.

90 Nowicki, S., Bonner, J., and Feather, B. Effects of locus of control and differential analogue interview procedures on the perceived therapeutic relationship. Journal of Consulting and Clinical Psychology, 1972, 38, 434-438.

91 Nowicki, S. and Duke, M. A preschool and primary internal-external control scale. Developmental Psychology, 1974, 10, 874-880.

92 Nowicki, S. and Segal, W. Perceived parental characteristics, locus of control orientation, and behavioral correlates of locus of control. Developmental Psychology, 1974, 10, 33-37.

93 Nowicki, S. and Strickland, B. A locus of control scale for children. Journal of Consulting and Clinical Psychology, 1973, 40, 148-154.

94a,b O'Leary, M., Rohsenow, D., and Donovan, D. Locus of control and patient attrition from an alcoholism treatment program. Journal of Consulting and Clinical Psychology, 1976, 44, 686-687.

95 Organ, D. Extraversion, locus of control. and individual differences in conditionability in organizations. Journal of Applied Psychology, 1975, 60, 401-404.

96 Parent, J., Forward, J., Canter, R., and Mohling, J. Interactive effects of teaching strategy and personal locus of control on studert performance and satisfaction. Journal of Educational Psychology, 1975, 67, 764-769.

97 Phares, E. Internal-external control as a determinant of amount of social influence exerted. Journal of Personality and Social Psychology, 1965, 2, 642-647.

98 Phares, E. Internal-external control and the reduction of reinforcement value after failure. Journal of Consulting and Clinical Psychology, 1971, 37, 386-390.

99 Phares, E. and Lamiell, J. Relationship of internal-external control of defensive preferences. Journal of Consulting and Clinical Psychology, 1974, 42, 872-878.

100 Phares, E., Ritchie, D., and Davis, W. Internal-external control and reaction to threat. Journal of Personality and Social Psychology, 1968, 10, 402-405.

101 Phares, E., Wilson, K., and Klyver, N. Internal-external control and the attribution of blame under neutral and distractive conditions. Journal of Personality and Social Psychology, 1971, 18, 285-288.

102 Pines, H. An attributional analysis of locus of control orientation and source of informational dependence. Journal of Personality and Social Psychology, 1973, 26, 262-272.

103 Prociuk, T. and Breen, L. Defensive externality and its relation to academic performance. Journal of Personality and Social Psychology, 1975, 31, 549-556.

104 Riemer, B. Influence of causal beliefs on affect and expectancy. Journal of Personality and Social Psychology, 1975, 31, 1163-1167.

105 Rotter, J. and Mulry, R. Internal versus external control of reinforcement and decision time. Journal of Personality and Social Psychology, 1965, 2, 598-604.

106 Ryckman, R., Gold, J., and Rodda, W. Confidence rating shifts and performance as a function of locus of control, self-esteem, and initial task experience. Journal of Personality and Social Psychology, 1971, 18, 305-310.

107 Ryckman, R. and Rodda, W. Locus of control and initial task experience as determinants of confidence changes in a chance situation. Journal of Personality and Social Psychology, 1971, 18, 116-119.

108 Runyon, K. Some interactions between personality variables and management styles. Journal of Applied Psychology, 1973, 57, 288-294.

109 Schneider, J. Skill versus chance activity preference and locus of control. Journal of Consulting and Clinical Psychology, 1968, 32, 333-337.

110 Schwartz, S. Awareness of interpersonal consequences, responsibility denial, and volunteering. Journal of Personality and Social Psychology, 1974, 30, 57-63.

111 Sherman, S. Internal-external control and its relationship to attitude change under different social influence techniques. Journal of Personality and Social Psychology, 1973, 26, 23-29.

112 Smith, I. Impact of computer-assisted instruction on student attitudes. Journal of Educational Psychology, 1973, 64, 366-372.

113 Smith, R. Changes in locus of control as a function of life crisis resolution. Journal of Abnormal Psychology, 1970, 75, 328-332.

114 Snyder, C. and Larson, G. A further look at student acceptance of general personality interpretations. Journal of Consulting and Clinical Psychology, 1972, 38, 384-388.

115 Sosis, R. Internal-external control and the perception of responsibility of another for an accident. Journal of Personality and Social Psychology, 1974, 30, 393-399.

116 Stone, G. and Jackson, T. Internal-external control as a determinant of the effectiveness of modeling and instructions. Journal of Counseling Psychology, 1975, 22, 294-298.

117 Straits, B. and Sechrest, L. Further support of some findings about the characteristics of smokers and nonsmokers. Journal of Consulting Psychology, 1963, 27, 282.

118 Strassberg, D. Relationships among locus of control, anxiety, and valued-goal expectations. Journal of Consulting and Clinical Psychology, 1973, 41, 319.

119 Strassberg, D. and Robinson, J. Relationship between locus of control and other personality measures in drug users. Journal of Consulting and Clinical Psychology, 1974, 42, 744-745.

120 Sundberg, N., Rohila, P., and Tyler, L. Values of Indian and American adolesecents. Journal of Personality and Social Psychology, 1970, 16, 374-397.

121 Tobacyk, J., Broughton, A., and Vaught, G. Effects of congruence-noncongruence between locus of control and field dependence on personality functioning. Journal of Consulting and Clinical Psychology, 1975, 43, 81-85.

122 Tolor, A. and Jalowiec, J. Body boundary, parental attitudes, and internal-external expectancy. Journal of Consulting and Clinical Psychology, 1968, 32, 206-209.

123 Tolor, A. and Reznikoff, M. Relation between insight, repression-sensitization, internal-external control, and death anxiety. Journal of Abnormal Psychology, 1967, 72, 426-430.

124 Tseng, M. Locus of control as a determinant of job proficiency, employability, and training satisfaction of vocational rehabilitation clients. Journal of Counseling Psychology, 1970, 17, 487-491.

125 Walls, R. and Miller, J. Delay of gratification in welfare and rehabilitation clients. Journal of Counseling Psychology, 1970, 17, 383-384.

126a,b Wallston, B., Wallston, K., Kaplan, G., and Maides, S. Development of the Health Locus of Control (HLC) Scale. Journal of Consulting and Clinical Psychology, 1976, 44, 580-585.

127 Warehime, R. and Foulds, M. Perceived locus of control and personal adjustment. Journal of Consulting and Clinical Psychology, 1971, 37, 250-252.

128 Wener, A. and Rehm, L. Depressive affect: A test of behavioral hypotheses. Journal of Abnormal Psychology, 1975, 84, 221-227.

129 Weiner, B., Heckhausen, H., Meyer, W., and Cook, R. Causal ascriptions and achievement behavior: A conceptual analysis of effort and reanalysis of locus of control. Journal of Personality and Social Psychology, 1972, 21, 239-248.

130 Weiner, B. and Kukla, A. An attributional analysis of achievement motivation. Journal of Personality and Social Psychology, 1970, 15, 1-20.

131 Weiner, B. and Potepan, P. Personality characteristics and affective reactions toward exams of superior and failing college students. Journal of Educational Psychology, 1970, 61, 144-151.

132 Weiss, H. and Sherman, J. Internal-external control as a predictor of task effort and satisfaction subsequent to failure. Journal of Applied Psychology, 1973, 57, 132-136.

133 Williams, C. and Nickels, J. Internal-external control dimension as related to accident and suicide proneness. Journal of Consulting and Clinical Psychology, 1969, 33, 485-494.

134 Williams, J. and Stack, J. Internal-external control as a situational variable in determining information seeking by Negro students. Journal of Consulting and Clinical Psychology, 1972, 39, 187-193.

135 Wolfe, R. Perceived locus of control and prediction of own academic performance. Journal of Consulting and Clinical Psychology, 1972, 38, 80-83.

136 Wolk, S. Situational constraint as a moderator of the locus of control-adjustment relationship. Journal of Consulting and Clinical Psychology, 1976, 44, 420-427.

137 Zytkoskee, A., Strickland, B., and Watson, J. Delay of gratification and internal versus external control among adolescents of low socioeconomic status. Developmental Psychology, 1971, 4, 93-98.

APPENDIX B

QUALITY OF STUDY RATING FORM
AND
FORM FOR THE EVALUATION OF AN ARTICLE

Title of article_____

Name of researcher _____

Quality of Study Rating Form

Criteria	Totally Unacceptable(1)	Poor (2)	Mediocre (3)	Good (4)	Excellent (5)
1. The research problem or purpose of the study is specified, and is sufficiently delimited so as to be amenable to investigation.					
2. Hypotheses are testible, are stated in such a manner as to be subject to empirical verification.					
3. Hypotheses either have an adequate theoretical rationale, presented in the article, or are logically derived from specified earlier research.					
4. Assumptions that form a basis for the study are specified and discussed.					
5. Method of sampling and assignment procedures are either specified or described.					
6. Sampling and assignment procedures are adequate for the purposes of the study					
7. Data-gathering methods or procedures are appropriate for testing the hypotheses of the study.					
8. Validity and reliability data are reported for all instruments utilized.					
9. Instuments utilized are adequately valid and reliable.					
10. Research design is appropriate for testing the hypotheses of the study.					
11. Research design is free of specific weaknesses.					
12. Appropriate methods are selected to analyze the data.					
13. Conclusions are substantiated by the evidence presented.					
14. Generalizations regarding present findings are confined to the population from which the sample was drawn.					
15. Tone of the report displays an unbiased, impartial scientific attitude, and both positive and negative results are presented and discussed.					
16. Limitations and weaknesses of the study are specified and discussed.					
Rater's determination of the overall level of quality of the study.					

See page 156 of the Handbook in Research and Evaluation, by Stephen Isaac and William B. Michael, 1971, for Edwin Wandt's Form for the Evaluation of an Article

APPENDIX C

TABLES

TABLE A

W COEFFICIENTS FOR RANKINGS OF ITEM-BY-ITEM SCORING CRITERIA

ITEM 1

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3
3	3	3	3
1	1	1	1
4	4	4	4
2	2	2	2
5	5	5	5

100 % agreement, W = 1.0

ITEM 2

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3
3	3	3	3
5	5	5	5
4	4	4	4
2	2	2	2
1	1	1	1

100 % agreement, W = 1.0

ITEM 3

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3
1	1	1	1
5	5	5	5
4	4	4	4
3	3	3	3
2	2	2	2

100 % agreement, W = 1.0

ITEM 4

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3	Σ	Σ^2
3	3	3	4	13	169
2	2	2	3	9	81
4	4	4	2	14	196
5	5	5	5	20	400
1	1	1	1	4	16
				60	862

$$W = \frac{12s}{k^2(n^3-n)} = 12\,\frac{(862 - \frac{3600}{5})}{16(125-5)} = \frac{12\,(862-720)}{16\,(120)} = \frac{12\,(142)}{16(120)} = \frac{1704}{1920} = .887$$

TABLE A (CONT)

ITEM 5

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3	Σ	Σ^2
5	5	5	5	20	400
2	2	3	2	9	81
3	3	2	3	11	121
1	1	1	1	4	16
4	4	4	4	16	256
				60	874

$$W = \frac{12(874-720)}{16\ (120)} = \frac{1848}{1920} = .962$$

ITEM 6

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3	Σ	Σ^2
2	2	2	2	8	64
4	3	4	4	15	225
1	1	1	1	4	16
5	5	5	5	20	400
3	4	3	3	13	169
				60	874

$$W = \frac{12(874-720)}{16\ (120)} = \frac{1848}{1920} = .962$$

ITEM 7

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3
3	3	3	3
5	5	5	5
2	2	2	2
1	1	1	1
4	4	4	4

100 % agreement, W = 1.0

ITEM 8

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3	Σ	Σ^2
3	2	3	3	11	121
5	5	5	5	20	400
2	3	2	2	9	81
1	1	1	1	4	16
4	4	4	4	16	256
				60	874

$$W = \frac{12(874-720)}{16\ (120)} = \frac{1848}{1920} = .962$$

TABLE A (CONT)

ITEM 9

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3	Σ	Σ^2
4	4	4	4	16	256
2	3	3	2	10	100
3	2	2	3	10	100
5	5	5	5	20	400
1	1	1	1	4	16
				60	872

$$W = \frac{12(872-720)}{16(120)} = \frac{12(152)}{16(120)} = \frac{1824}{1920} = .950$$

ITEM 10

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3
3	3	3	3
5	5	5	5
4	4	4	4
2	2	2	2
1	1	1	1

100 % agreement, W= 1.0

ITEM 11

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3	Σ	Σ^2
3	3	3	3	12	144
4	5	4	4	17	289
2	2	2	1	7	49
5	4	5	5	19	361
1	1	1	2	5	25
				60	868

$$W = \frac{12(868-720)}{16(120)} = \frac{12(148)}{16(120)} = \frac{1776}{1920} = .925$$

ITEM 12

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3
3	3	3	3
2	2	2	2
4	4	4	4
5	5	5	5
1	1	1	1

100 % agreement, W =1.0

TABLE A (CONT)

ITEM 13

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3
1	1	1	1
3	3	3	3
2	2	2	2
5	5	5	5
4	4	4	4

100 % agreement, W = 1.0

ITEM 14

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3
1	1	1	1
2	2	2	2
4	4	4	4
3	3	3	3
5	5	5	5

100 % agreement, W = 1.0

ITEM 15

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3	Σ	Σ^2
3	2	4	3	12	144
1	1	3	1	6	36
5	5	5	5	20	400
2	3	1	2	8	64
4	4	2	4	14	196
				60	840

$$W = \frac{12(840-820)}{16(120)} = \frac{12(120)}{16(120)} = \frac{1440}{1920} = .750$$

ITEM 16

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3	Σ	Σ^2
4	4	2	5	15	225
2	2	4	2	10	100
3	3	3	3	12	144
1	1	1	1	4	16
5	5	5	4	19	361
				60	846

$$W = \frac{12(846-720)}{16(120)} = \frac{12(126)}{16(120)} = \frac{1512}{1920} = .787$$

RELIABILITY OF SCORING CRITERIA RANKINGS

$$\text{Mean } W = \frac{1.0+1.0+1.0+.887+.962+.962+1.0+.962+.95+1.0+.925+1.0+1.0+1.0+.75+.789}{16} = .949$$

TABLE B

W COEFFICIENTS FOR RANKINGS OF ITEM-BY-ITEM SCORING CRITERIA FOR THE EIGHT ALTERED ITEMS

ITEM 1

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3
3	3	3	3
1	1	1	1
4	4	4	4
2	2	2	2
5	5	5	5

100 % agreement, W -1.0

ITEM 2

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3
3	3	3	3
5	5	5	5
4	4	4	4
2	2	2	2
1	1	1	1

100 % agreement, W =1.0

ITEM 6

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3	Σ	Σ^2
2	2	2	2	8	64
4	3	3	3	13	169
1	1	1	1	4	16
5	5	5	5	20	400
3	4	4	4	15	225
				60	874

$$W = \frac{12s}{k^2(n^3-n)} = \frac{12(874-720)}{16(120)} = .962$$

ITEM 8

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3	Σ	Σ^2
3	3	2	2	10	100
5	5	5	5	20	400
2	2	3	3	10	100
1	1	1	1	4	16
4	4	4	4	16	256
				60	872

$$W = \frac{12(872-720)}{16(120)} = \frac{1824}{1920} = .950$$

TABLE B (CONT)

ITEM 9

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3
4	4	4	4
2	2	2	2
3	3	3	3
5	5	5	5
1	1	1	1

100% agreement, W = 1.0

ITEM 11

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3	Σ	Σ^2
3	3	3	4	13	169
4	5	4	5	18	324
2	2	2	2	8	64
5	4	5	3	17	289
1	1	1	1	4	16
				60	862

$$W = \frac{12(862-720)}{16(120)} = \frac{1704}{1920} = .887$$

ITEM 12

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3
3	3	3	3
2	2	2	2
4	4	4	4
5	5	5	5
1	1	1	1

100 % agreement, W= 1.0

ITEM 16

RESEARCHER	RELIABILITY ASST 1	RELIABILITY ASST 2	RELIABILITY ASST 3	Σ	Σ^2
4	4	4	4	16	256
2	3	2	2	9	81
3	2	3	3	11	121
1	1	1	1	4	16
5	5	5	5	20	400
				60	874

$$W = \frac{12(874-720)}{16(120)} = \frac{12(154)}{16(120)} = \frac{1848}{1920} = .962$$

RELIABILITY OF SCORING CRITERIA RANKINGS

$$\text{Mean } W = \frac{1.0+1.0+.962+.95+1.0+.887+1.0+.962}{8} = .970$$

TABLE C

OVERALL QUALITY OF STUDY RATINGS

STUDY	RESEARCHER	ASST 1	ASST 2	ASST 3	μ_i
1	3	2	3	3	2.75
2	4	4	4	3	3.75
3	1	1	2	1	1.25
4	3	1	3	3	2.50
5a	2	2	2	3	2.25
5b	1	1	2	1	1.25
6	3	4	2	3	3.00
7	4	4	3	4	3.75
8	4	4	4	3	3.75
9	1	1	1	1	1.00
10	1	1	1	1	1.00
11	3	4	2	3	3.00
12	1	1	1	1	1.00
13	3	3	2	4	3.00
14	1	1	2	1	1.25
15	4	4	4	3	3.75
16	4	4	4	4	4.00
17	3	1	1	4	2.25
18	4	4	4	4	4.00
19	4	4	3	4	3.75
20	1	1	2	3	1.75
21	5	4	4	4	4.25
22	2	3	2	3	2.50
23a	2	2	2	2	2.00
23b	3	3	2	4	3.00
24	3	4	3	3	3.25
25	4	4	3	4	3.75

RELIABILITY OF OVERALL RATINGS

$$r_{11} = \frac{ms_B - ms_W}{ms_B + (n-1)\, ms_W}$$

$$ms_B = \frac{n \Sigma (\mu i - u)^2}{s-1}$$

$$ms_B = \frac{4\ (29.9151)}{26} = \frac{119.6604}{26} = 4.6023$$

$$ms_W = \frac{\Sigma \Sigma\ (x - \mu i\ )^2}{N-s}$$

$$ms_W = \frac{31.25}{81} = .3858$$

$$r_{11} = \frac{ms_B - ms_W}{ms_B - (n-1)\, ms_W} = \frac{4.6023 - .3858}{4.6023 + 3(.3858)}$$

$$= \frac{4.2165}{4.6023 + 3.(.3858)} = \frac{4.2165}{5.7597} = .7321$$

TABLE D

ASSIGNMENTS TO CONTENT CATEGORIES

STUDY	RESEARCHER	RELIABILITY ASSISTANT
1	Personality	Personality
2	Attribution of Causality	Self Control of Behavior
3	Psychological Adjustment	Psychological Adjustment
4	Strategy Preferences & Learning	Academic Achievement
5	Psychological Adjustment	Psychological Adjustment
6	Interpersonal Percerption & Processes	Strategy Preferences & Learning
7	Alcoholism & Drug Abuse	Alcoholism & Drug Abuse
8	Information Acquisition & Use	Informantion Acquisition & use
9	Self Control of Behavior	Self Control of Behavior
10	Information Acquisition & Use	Information Acquisition & Use
11	Strategy Preferences & Learning	Personality
12	Psychological Adjustment	Psychological Adjustment
13	Interpersonal Perception & Processes	Personality
14	Social Influence Processes	Social Influence Processes
15	Psychological Adjustment	Personality
16	Interpersonal Perception & Processes	Job Competence & Satisfaction
17	Strategy Preferences & Learning	Strategy Preferences & Learning
18	Job Competence & Satisfaction	Job Competence & Satisfaction
19	Research on the I-E Scale	Research on the I-E Scale
20	Social Influence Processes	Social Influence Processes
21	Behavior & Individual or Group Counseling	Behavior & Individual or Group Counseling
22	Personality	Personality
23	LC Measurement	LC Measurement
24	Group Differences	Group Differences
25	Group Differences	Group Differences

RELIABILITY OF CONTENT CATEGORIES

P_o = .04+.04+.08+.04+.12+.04+.04+.04+.04+.08+.04+.08+.08 = .76

P_c = .0016+.0016+0096+.0016+.0192+.0016+.0032+.0096+.0032+.0064+ 0016+.0064+.016 = .0816

$K = \frac{.76-.0816}{1-.0816} = \frac{.6784}{.9184} = .7387$

TABLE E

OVERALL RATINGS, MEAN RATINGS, AND CATEGORY PLACEMENTS BY NUMBER OF STUDY

STUDY	OVERALL RATING	MEAN RATING	SOURCE OF DERIVATION OF HYPOTHESES	MEASUREMENT SCALE SELECTED	DESIGN OF STUDY	I VS. E DETERMINATION	FINDINGS POS	FINDINGS NEG
1	4	4.07	LC concepts	Rotter, Mirel's factors	True Exper.	Dichotomy	2	
2	1	3.27	Prev. research alone.	Rotter	Ex Post Facto	Dichotomy	3	
3	3	2.69	LC concepts	Rotter	Ex Post Facto	Not reported	1	1
4	3	3.60	LC concepts	Rotter	Quasi-Exper.	Trichotomy	1	1
5	4	3.52	LC concepts	IAR	Quasi-Exper.	Dichotomy	2	
6	3	2.79	LC concepts	Rotter	Quasi-Exper.	Not Applicable(NA)	1	
7	3	3.62	LC concepts	Rotter	Correlational	NA		1
8	3	3.08	LC concepts	Rotter	Ex Post Facto	NA	2	
9	3	3.13	LC concepts	Rotter	True Exper.	Dichotomy		2
10	3	3.21	LC concepts	Rotter	Correlational	NA		2
11	3	3.37	Prev. research alone	James	True Exper.	Extreme groups	1	1
12	4	3.64	LC concepts	Rotter	True Exper.	Dichotomy	2	5
13	3	3.27	LC concepts	Rotter	Correlational	NA	4	
14	1	3.67	LC theory	Nowicki-Strickland (Adults)	Ex Post Facto	Extreme groups	2	
15	4	3.62	LC concepts	Levenson	Ex Post Facto	Dichotomy	2	
16	1	2.60	LC concepts	Levenson	Ex Post Facto	NA	2	2
17	1	3.00	LC concepts	Rotter	Quasi-Exper.	Extreme Scores	1	
18	3	3.21	LC concepts	Rotter, Mirel's Factors	True Exper.	Extreme Groups	1	

TABLE E (CONT)

STUDY	OVERALL RATING	MEAN RATING	SOURCE OF DERIVATION OF HYPOTHESES	MEASUREMENT SCALE SELECTED	DESIGN OF STUDY	I VS. E DETEMINATION	FINDINGS POS.	FINDINGS NEG
19	4	3.56	Could not be determined	Rotter	True Exper.	NA		1
20 a	2	3.14	Prev.research alone	Rotter	Correlational	NA		1
20b	2	3.14	Prev. research alone	Rotter	Correlational	NA		1
21	4	3.21	Did not fit categories	Rotter	Correlational	NA	1	
22	4	3.57	Prev. research alone	Rotter	Correlational	NA	1	
23 a	4	3.64	LC concepts	Rotter	True Exper.	NA	1	
23b	4	3.64	LC concepts	Rotter	True Exper.	NA		1
24	3	3.38	LC concepts	Rotter	Ex Post facto	Dichotomy	3	1
25	1	2.77	Prev. research alone	Rotter	Quasi-Exper.	NA	1	
26	1	3.64	LC concepts	Rotter	Ex Post Facto	Dichotomy	1	
27	3	3.46	Prev. research alone	Nowicki-Strickland (Adults)	Ex Post Facto	NA	2	
28 a	1	2.87	Did not fit categories	ALOE (Miller)	Ex Post Facto	NA	1	
28b	1	2.87	Did not fit categories	ALOE (Miller)	Ex Post Facto	NA	1	
29	5	3.50	LC concepts	IAR	Ex Post Facto	NA	3	1
30	3	3.00	LC concepts	Rotter	Quasi-Exper.	Not-reported	3	
31	1	3.27	Did not fit categories	Rotter	Ex Post Facto	Dichotomy	3	
32	3	3.53	LC concepts for Hyp 1 LC concepts for Hyp 2	IAR	True Exper.	NA	2	
33	3	3.53	LC concepts	Rotter	Quasi-Exper.	NA	2	2
34	3	2.85	Prev. research alone	Rotter	Quasi-Exper.	Extreme groups	1	
35	3	2.85	LC theory	Researcher's own	Ex Post Facto	NA	1	1
36	1	2.71	Did not fit categories	Rotter	Ex Post Facto	NA	1	
37	1	3.64	Did not fit categories	Rotter	Quasi-Exper.	Dichotomy	1	

TABLE E (CONT)

STUDY	OVERALL RATING	MEAN RATING	SOURCE OF DERIVATION OF HYPOTHESES	MEASUREMENT SCALE SELECTED	DESIGN OF STUDY	I VS. E DETERMINATION	FINDINGS POS	FINDINGS NEG
38	3	2.79	LC theory	IAR	Quasi-Exper.	Extreme groups	1	2
39a	2	3.80	LC theory	Rotter	Ex Post Facto	NA	1	
39b	1	3.27	LC theory	Levenson	Ex Post Facto	NA	2	1
40	3	3.14	LC concepts	Rotter	Quasi- Exper.	Extreme groups	1	
41	3	3.07	LC concepts	Rotter	Ex Post Facto	NA	1	1
42	2	3.21	LC concepts	Rotter	Ex Post Facto	Trichotomy	1	1
43	4	3.54	Prev. research alone	Gruen, Korte,& Stephens I-E	Ex Post Facto	NA	1	
44	3	3.08	LC concepts	Rotter	Ex Post Facto	NA	1	1
45	3	3.00	LC concepts	Rotter	Correlational	Extreme groups	1	
46	4	4.14	LC concepts	Rotter	True Exper.	Trichotomy		2
47	4	3.31	LC theory	James	True Exper.	Extreme groups	2	
48	4	3.54	LC concepts	Rotter	True Exper.	Trichotomy	2	
49	4	3.29	LC concepts	Rotter	True Exper.	Dichotomy	3	1
50	4	3.25	LC concepts	Rotter	True Exper.	Dichotomy	2	4
51	4	4.00	LC concepts	IRA	True Exper.	Trichotomy	1	1
52	1	3.46	LC concepts	Rotter, Mirel's factors	True Exper.	Trichotomy	3	4
53	4	3.31	LC concepts	Rotter	Ex Post Facto	NA	1	
54	1	3.00	Prev. research only	James	Ex Post Facto	NA	3	3
55	3	3.44	Did not fit categories	Researcher's own	Correlational	NA	1	1
56	3	2.58	Prev. research alone	IRA	Ex Post Facto	NA	1	1

TABLE E (CONT)

STUDY	OVERALL RATING	MEAN RATING	SOURCE OF DERIVATION OF HYPOTHESES	MEASUREMENT SCALE SELECTED	DESIGN OF STUDY	I VS. E DETERMINATION	FINDINGS POS	FINDINGS NEG
57	3	2.93	Prev. research alone	Rotter	Ex Post Facto	Trichotomy	1	1
58	3	3.27	LC concepts	Nowicki-Strickland (Adults)	True Exper.	Dichotomy	1	1
59	4	3.36	LC concepts	Rotter	True Exper.	Extreme groups	1	1
60	4	3.23	LC theory for Hyp1 LC concepts for Hyp2	Rotter	Quasi-Exper.	Dichotomy	2	
61a	4	3.38	LC theory	Rotter	Ex Post Facto	Trichotomy	1	
61b	3	2.92	LC theory	Rotter	Quasi-Exper.	Dichotomy		1
62	4	2.92	LC concepts	Rotter	Ex Post Facto	Dichotomy	1	1
63a	1	3.07	Prev. research alone	Rotter	True Exper.	Dichotomy	1	4
63b	1	3.00	Prev. research alone	Rotter	True Exper.	Dichotomy	3	3
64	1	3.21	LC concepts	Rotter	True Exper.	Extreme groups		3
65a	4	3.92	Prev. research alone	Researcher's own	Ex Post Facto	Dichotomy		1
65b	4	3.92	Prev. research alone	Researcher's own	Ex Post Facto	Dichotomy	1	
65c	4	3.92	Prev. research alone	Researcher's own	Ex Post Facto	Dichotomy		1
66	3	3.36	LC concepts	Rotter	Ex Post Facto	Trichotomy	1	1
67	1	2.92	LC theory	Researcher's own	Ex Post Facto	NA	1	1
68	1	3.47	LC concepts	Rotter	Quasi-Exper.	Dichotomy	4	
69	3	2.92	LC concepts	1962 Rotter	Ex Post Facto	NA	2	
70	3	3.14	Prev. research alone	Rotter	Ex Post Facto	Dichotomy	1	1
71	1	3.00	LC theory	Levenson	Ex Post Facto	NA	5	2
72a	3	3.71	LC theory	Levenson	Ex Post Facto	NA	1	
72b	4	3.79	LC theory	Levenson	Ex Post Facto	NA	1	
73	3	3.43	Did not fit categories	Rotter	Ex Post Facto	Dichotomy	4	1
74	3	3.47	Prev. research alone	James	Ex Post facto	NA	1	
75	1	3.29	Did not fit categories	Rotter	Correlational	NA	1	1

TABLE E (CONT)

STUDY	OVERALL RATING	MEAN RATING	SOURCE OF DERIVATION OF HYPOTHESES	MEASUREMENT SCALE SELECTED	DESIGN OF STUDY	I VS. E DETERMINATION	FINDINGS POS	FINDINGS NEG
76	4	3.43	LC concepts	James	Ex Post Facto	Extreme groups		1
77	1	3.29	Prev. research alone	Rotter	Ex Post Facto	NA	1	
78	3	3.36	LC concepts	Rotter	Correlational	NA	1	1
79	3	3.00	LC concepts	Rotter	Ex Post Facto	Extreme groups	2	2
80	3	3.21	LC concepts	Nowicki-Strickland (Adult)	Ex Post Facto	NA	6	1
81	1	2.93	LC theory	Rotter	Ex Post Facto	NA	1	
82a	3	2.86	Could not be determined	Rotter	Ex Post Facto	Extreme groups	1	
82b	3	2.71	Could not be determined	Rotter	Ex Post Facto	Extreme groups	1	
83	4	3.33	LC theory	Researcher's own	Ex Post Facto	NA	1	2
84a	3	3.33	LC concepts	Attribution measure	True Exper.	Dichotomy		1
84b	4	3.42	LC Theory	Attribution measure	Quasi-Exper.	Dichotomy		1
85	4	4.00	LC theory	Rotter	True Exper.	Dichotomy		1
86	2	3.21	LC concepts	Stanford Preschool I-E	Correlational	NA	3	
87	1	3.43	LC concepts	Rotter	Ex Post Facto	NA		1
88	3	3.36	Prev.research alone	Rotter	Correlational	NA	4	
89	4	3.79	Prev. research alone	Rotter	True Exper.	Trichotomy	1	1
90	4	3.86	Prev. research alone	Rotter	True Exper.	Not reported	1	1
91	4	3.87	Prev. research alone	Preschool & Primary Nowicki-Strickland	Correlational	NA	1	1
92	3	3.27	Prev. research alone	Nowicki-strickland (both child & Adult form)	Correlational	NA	2	1
93	4	3.54	Could not be determined	Nowicki-Strickland (child)	Correlational	NA	2	1
94a	3	3.00	Prev. research alone	Rotter, Mirels factors	Ex Post Facto	NA		1
94b	1	2.77	Pref. research alone	Rotter	Ex Post Facto	NA		1

TABLE E (CONT)

STUDY	OVERALL RATING	MEAN RATING	SOURCE OF DERIVATION OF HYPOTHESES	MEASUREMENT SCALE SELECTED	DESIGN OF STUDY	I VS. E DETERMINATION	FINDINGS POS	FINDINGS NEG
95	3	3.14	LC concepts	Rotter	Ex Post Facto	Dichotomy	1	1
96	2	2.92	LC concepts	Rotter,Mirel's Factors	True Exper.	Trichotomy	1	
97	3	2.62	LC concepts	Rotter	Quasi-Exper.	Extreme groups	1	
98	1	2.86	LC theory	Rotter	Ex Post Facto	Extreme groups	1	
99	2	3.36	LC concepts	Rotter	Ex Post Facto	Extreme groups	1	
100	2	3.57	LC concepts	Rotter	Ex Post Facto	Extreme groups	2	2
101	1	3.08	LC concepts	Rotter	Quasi-Exper.	Extreme groups	3	1
102	4	3.60	LC theory	1962 Rotter	True Exper.	Dichotomy	3	1
103	4	3.53	LC concepts	Levenson	Ex Post Facto	Did not fit categories	4	2
104	4	3.54	LC theory	Researcher's own	Quasi-Exper.	Did not fit categories	1	1
105	3	3.00	LC concepts	1962 Rotter	True Exper.	Dichotomy	1	1
106	3	3.80	LC theory	Rotter	Correlational	Trichotomy	1	
107	5	3.86	LC concepts	Rotter	True Exper.	Trichotomy	1	1
108	4	3.67	LC concepts	Rotter	True Exper.	Trichotomy	2	1
109	4	3.62	LC concepts	Rotter	Correlational	NA	2	2
110	3	3.31	Prev. research alone	Rotter	Correlational	NA	1	
111	1	2.86	LC concepts	Rotter	Quasi-Exper.	Extreme groups	2	
112	1	3.21	Prev. research alone	Crandall	Quasi-Exper.	NA		1
113	3	3.29	LC concepts	Rotter	Ex Post Facto	NA	1	2
114	3	2.77	LC concepts	Rotter	Correlational	NA	1	

TABLE E (CONT)

STUDY	OVERALL RATING	MEAN RATING	SOURCE OF DERIVATION OF HYPOTHESES	MEASUREMENT SCALE SELECTED	DESIGN OF STUDY	I VS. E DETERMINATION	FINDINGS POS	FINDINGS NEG
115	2	3.07	LC theory	Rotter	Ex Post Facto	Trichotomy	7	
116	5	4.07	Prev. research alone	Rotter, Mirel's factors	True Exper.	Extreme groups	1	1
117	1	2.71	LC concepts	James	Ex Post Facto	NA		1
118	4	3.08	LC theory	Rotter	Correlational	Trichotomy	2	
119	4	3.29	Prev. research alone	Rotter	Correlational	NA	1	1
120	1	3.36	Prev. research alone	Researcher's own	Ex Post Facto	NA	1	1
121	3	2.67	Prev. research alone	Rotter	Ex Post Facto	Extreme groups	2	
122	2	3.12	LC concepts	Rotter	Correlational	NA	1	1
123	3	3.21	LC concepts	Rotter	Correlational	NA	2	
124	1	2.71	Prev. research alone	Rotter	Ex Post Facto	Extreme groups	4	5
125	1	2.77	Prev. research alone	Rotter	Ex Post Facto	Dichotomy		1
126a	2	3.60	LC theory	Rotter, Health LC Scale	Ex Post Facto	Dichotomy	1	1
126b	3	3.38	LC theory	Rotter, Health LC Scale	True Exper.	Not reported		1
127	3	3.20	Prev. research alone	Rotter	Correlational	NA	1	1
128	3	3.21	LC theory	Rotter	True Exper.	Dichotomy		1
129	3	3.00	LC theory	Researcher's own	Ex Post Facto	NA		1
130	2	3.00	LC concepts	IAR	Ex Post Facto	NA	1	1
131	1	2.66	Prev. research alone	IAR	Ex Post Facto	NA	2	3
132	3	3.00	LC theory for Hyp 1 LC concepts for Hyps 2&3	Rotter	Correlational	NA	2(H_1,H_2)	1 (H_3)
133	4	3.64	Prev. research alone	Rotter	Correlational	Trichotomy	2	2
134	3	3.47	LC theory	Rotter	True Exper.	Dichotomy	2	2

TABLE E (CONT)

STUDY	OVERALL RATING	MEAN RATING	SOURCE OF DERIVATION OF HYPOTHESES	MEASUREMENT SCALE SELECTED	DESIGN OF STUDY	I VS. E DETERMINATION	FINDINGS POS.	FINDINGS NEG
135	2	3.14	LC concepts	Rotter	Ex Post Facto	Trichotomy	1	2
136	2	3.57	LC theory	Nowicki-Strickland	Ex Post Facto	NA	2	
137	4	3.15	LC concepts	Bialer	Quasi-Exper.	NA	1	1

APPENDIX D

ITEMS AND SCORING CRITERIA FOR QUALITY OF STUDY RATING FORM

ITEMS AND SCORING CRITERIA FOR QUALITY OF STUDY RATING FORM

1. The research problem or purpose of the study is specified, and is sufficiently delimited as to be amenable to investigation.

Scoring Points for the Five Quality of Study Rating Levels for Item 1:

__5__ Research problem or purpose of study is clearly identified by the language chosen ("the problem of the study was..." or " the purpose of the study was..." etc.), and would appear delimited enough to be amenable to investigation.

__4__ Research problem or purpose of study is not clearly identified by the language chosen, but the inference that a statement equivalent to one of these was intended is fairly easily made, and the problem would appear delimited enough to be amenable to investigation.

__3__ Reader has to work pretty hard to figure out the research problem or purpose of the study, but the problem would appear to be amenable to investigation.

__2__ It is questionable whether problem is amenable to investigation.

__1__ Problem is not amenable to investigation.

2. Hypotheses are testable, are stated in such a manner as to be subject to empirical verification.

Scoring Points for the Five Quality of Study Rating Levels for Item 2:

__5__ Subjects' statuses on all independent and dependent variables mentioned in the hypotheses are capable of being measured, and hypotheses are clearly identified by the language chosen.

__4__ Subjects' statuses on all independent and dependent variables mentioned in the hypotheses are capable of being measured and the language chosen makes identification of the hypotheses fairly easy.

__3__ Subjects' statuses on all independent and dependent variables mentioned in the hypotheses are capable of being measured, but the reader has to work pretty hard to identify the hypotheses of the study.

__2__ It is questionable whether subjects statuses on all independent or dependent variables mentioned in the hypotheses are capable of being measured.

__1__ Subjects' statuses on one or more of the independent or dependent variables mentioned in the hypotheses are not capable of being measured.

3. Hypotheses either have an adequate theoretical rationale, presented in the article, or are logically derived from specified earlier research.

Scoring Points for the Five Quality of Study Rating Levels for Item 3:

__5__ Hypotheses are derived from a clearly presented theoretical rationale, or from specified earlier research, and no significant gaps in logic are apparent in the derivations.

__4__ Derivation of the hypotheses is lacking somewhat in clarity, but no significant gaps in logic are apparent in the derivations.

__3__ Some significant gaps in logic are apparent in the derivation of the hypotheses, but is is reasonably easy to fill them in with one's own assumptions and inferences.

__2__ It is difficult to determine where the hypotheses came from, but it is possible to reconstruct some kind of derivation for them.

__1__ Derivation of the hypotheses cannot be determined.

4. Researcher's assumptions are specified and discussed. (Do not rate on this item if simply cannot determine whether assumptions were, or should have been, presented).

Scoring Points for the Five Quality of Study Rating Levels for Item 4:

_5_ Assumptions involved in the development of the research problem or the derivation of the hypotheses are specified and discussed.

_4_ Assumptions not directly connected with the development of the research problem or the derivation of the hypotheses are specified and discussed.

_3_ Researcher neglects to present and discuss one or more assumptions that are not directly connected with the development of the research problem or the derivation of the hypotheses.

_2_ Researcher makes some reference to an assumption involved in the development of the research problem or the derivation of the hypotheses, but he fails to clarify how it fits into the discussion.

_1_ Researcher neglects to present and discuss assumptions involved in the development of the problem or the derivation of the hypotheses.

5. Method of sampling and assignment procedures are either specified or described. (Correlational studies will not generally involve group assignments, and should not, therfore, be penalized for not making reference to assignment procedures.)

Scoring Points for the Five Quality of Study Rating Levels for Item 5:

_5_ Method of sampling and assignment procedures are clearly identified by the language chosen.

_4_ Method of sampling and assignment procedures are not clearly identified by the language chosen, but it is fairly easy to infer what was done.

_3_ Reader has to work pretty hard to figure out what method of sampling and assignment procedures consisted of.

_2_ Information provided is sufficient to figure out what was done with respect to sampling, but not assignment, or with respect to assignment, but not sampling.

_1_ Neither method of sampling nor assignment procedures can be determined.

6. Sampling and assignment procedures are adequate for the purposes of the study. (Do not rate here if could not determine, in the case of ex post facto studies, what method of smapling consisted of, or, in the case of experimental studies, what both method of sampling and assignment procedures consisted of .)

Scoring Points for the Five Quality of Study Rating Levels for Item 6:

_5_ Experimental research only: Subjects were randomly selected, then randomly assigned to treatment groups. Ex post facto research only : Subjects were randomly selected.

_4_ Experimental research only: Lack of random sampling, but subjects were randomly assigned to treatment groups.

_3_ Ex post facto research only : Lack of random sampling, but, procedures utilized would not appear to have introduced significant selection biases. Experimental research only: Lack of both random sampling and random assignment but other reasonably adequate procedures for insuring the initial equation of experimental and control groups were utilized.

_2_ Ex post facto research only: Sample was composed of volunteers.

_1_ Experimental research only: Lack of both random sampling and random assignment, in the absence of other adequate procedures for insuring the initial equation of experimental and control groups.

7. Data-gathering methods or procedures are appropriate for testing the hypotheses of the study.

Scoring Points for the Five Quality of Study Rating Levels for Item 7:

__5__ The methods or procedures selected appear ideally suited for obtaining the data needed.

__4__ The methods or procedures selected may not be the best available, but they can do a good job of obtaining the data needed.

__3__ The methods or procedures selected are adequate, nothing more.

__2__ It is questionable whether the data needed can be obtained through the methods or procedures selected.

__1__ The data needed cannot be obtained through the methods or procedures selected.

8. Validity and reliability data are reported for all instruments utilized.

Scoring Points for the Five Quality of Study Rating Levels for Item 8:

__5__ Reliability coefficients are reported, and information regarding validity provided, for all instruments utilized.

__4__ Reliability coefficients are reported for all instruments utilized. Information regarding validity may or may not be provided.

__3__ Reliability coefficients are reported for some, but not all instruments utilized, or summations of reliability findings are provided. Information regarding validity may or may not be provided.

__2__ Researcher neither reports reliability coefficients, nor provides summations of reliability findings, but he does specifically tell the reader where to locate them. (An unexplained reference following an identification of the instrument is <u>insufficient</u>; the researcher must specifically tell the reader that validity and/or reliability data may be found in the reference in question) <u>or</u>, researcher provides information regarding validity findings only.

__1__ Researcher neither reports nor provides summations of reliability findings, nor does he specifically tell the reader where to locate them himself.

9. Instruments utilized are adequately valid and reliable. (Do <u>not</u> rate on this item when reliability information was not provided.)

Scoring Points for the Five Quality of Study Rating Levels for Item 9:

__5__ All reliability coefficients reported were in the .80's or above. Such validity findings as were reported were generally supportive of the validity of the instrument.

__4__ All reliability coefficients reported were in the 70's or above. Such validity findings as were reported were generally supportive of the validity of the instrument.

__3__ All reliability coefficients reported were in the 60's or above, <u>or</u>, reliability findings were described as favorable, though no reliability coefficients were provided. Such validity findings as were reported were generally supportive of the validity of the instrument.

__2__ Summaries of validity and/or reliability studies indicated that findings were equivocal.

__1__ Some of the reliability coefficients reported were in the .50's or below, <u>or</u>, summaries of validity and/or reliability studies indicated that findings were generally negative.

10. Research design is appropriate for testing the hypotherses of the study

Scoring Points for the Five Quality of Study Rating Levels for Item 10:

__5__ The research design chosen appears ideally suited for obtaining the data needed.

__4__ The research design chosen may not be the best available, but it can do a good job of obtaining the data needed.

__3__ The research design chosen is adequate, nothing more.

__2__ It is questionable whether the data needed can be obtained by the research design chosen.

__1__ The data needed cannot be obtained by the research design chosen.

11. Research design is free of specific weaknesses. (Refer to Campbell and Stanley handout, attached, for explication of the biases referred to, below.)

Scoring Points for the Five Quality of Study Rating Levels for Item 11:

__5__ None of the validity threats identified by Campbell and Stanley are present, or, only difficulties related to the reactive effects of the experimental arrangements or multiple treatment interference is present.

__4__ Or, experimental mortality difficulties are present, but there is evidence that they were adequately controlled for, biases resulting from pre-testing history, or maturation are present, but a control group was used, or, for experimental research only, potential selection biases due to lack of random sampling are present, but these were controlled for by the researcher's randomly assigning subjects to treatment groups*

__3__ Biases due to instrumentation or experimental mortality difficulties are present, or, for ex post facto research only, where random sampling was not employed, procedures that were utilized would not appear to have introduced significant selection biases, or, for experimental research only, potential selection biases due to lack of both random sampling and random assignment are present, but other reasonably adequate procedures for ensuring the initial equation of experimental and control groups were applied.

__2__ Biases resulting from pre-testing, history, or maturation are present and, no control group was used, or for ex post facto research only, potential selection biases due to reliance on volunteer subjects are present.

__1__ For experimental research only: Selection biases due to lack of random sampling are present, and, it is clear that there was a lack of random assignment to treatment groups, plus an absence of other adequate procedures for insuring the initial equation of experimental and control groups.

12. Appropriate methods are selected to analyze the data.

Scoring Points for the Five Quality of Study Rating Levels for Item 12:

__5__ Methods selected appear ideally suited for testing the relationships hypothesized, and, no unacceptable methods were used anywhere in the analysis.

__4__ Methods selected can do a good job of testing the relationships hypothesized. and, no unacceptable methods were used anywhere in the analysis.

__3__ Methods selected are adequate, nothing more, and, no unacceptable methods were used anywhere in the analysis.

__2__ It is questionable whether the methods selected can do an adequate job of testing the relationships hypothesized, or, unacceptable methods, (to include multiple t's, the Duncan Multiple Range Test, and multiple F's used to conduct pairwise comparisons) were used in the analysis. but were not used to analyze data involved in the tests of the hypotheses.**

__1__ It is clear that the methods selected cannot do an adequate job of testing the relationships hypothesized, or, unacceptable methods were used to analyze data involved in the tests of the hypotheses.

---

* In cases where it cannot be determined what method of sampling and/or assignment procedures consisted of, the researcher must be given the benefit of the doubt. Ex post facto and correlational studies should be treated in a like manner, with respect to discriminations between rankings of 3 and 2.

** Some qualification of this statement is necessary: The three methods referred to here are to be considered unacceptable whenever the significance levels claimed by the researcher multiplied by the number of pairwise comparisons made by one of these methods exceeds .10. As long as this product does not exceed .10 the methods are to be considered acceptable, and thus not penalized. Planned orthogonal comparisons utilizing t and F statistics are exempted from these guidelines as they do not result in unacceptable increases in alpha.

13. Conclusions are substantiated by the evidence presented. (A distinction should be drawn here between speculative statements and conclusions. This item is concerned only with conclusions.)

Scoring Points for the Five Quality of Study Rating Levels for Item 13:

__5__ It is clear that all conclusions made by the researcher are substantiated by the evidence presented.

__4__ It would appear that all conclusions made by the researcher are substantiated by the evidence presented.

__3__ It is questionable whether all conclusions made by the researcher are substantiated by the evidence presented, but it is difficult to decide one way or the other.

__2__ It would appear that one or more conclusions made by the researcher is not substantiated by the evidence presented.

__1__ It is clear that one or more conclusions made by the researcher is not substantiated by the evidence presented.

14. Generalizations regarding present findings are confined to the population from which the sample was drawn. (don't worry, here, about over-inclusive descriptions; penalize here only when present findings are generalized beyond the population from which they were obtained)

Scoring Points for the Five Quality of Study Rating Levels for Item 14:

__5__ Researcher stuck strictly to the population from which his sample was obtained, retaining all relevant qualifers regarding the sample and population: "The present findings indicate that externally-oriented undergraduates..."

__4__ Researcher extended beyond his population some (though he stayed within the immediately superordinate population), but, he did qualify his generalization with speculative language: "The present findngs suggest that externally-oriented college students..."

__3__ Researcher extended beyond the immediately superordinate population, but, he did qualify his generalization with speculative language: "The present findings suggest that externals..."

__2__ Researcher extended beyond his population some (though he stayed within the immediately superordinate population), and, failed, at the same time, to qualify his generalization with speculative language: "The present findings indicate that externally-oriented college students..."

__1__ Researcher extended beyond the immediately superordinate population, and, failed, at the same time, to qualify his generalization with speculative language: "The present findings indicate that externals.."

15. Tone of the report displays an unbiased, impartial scientific attitude, and both positive and negative results are presented and discussed.

Scoring Points for the Five Quality of Study Rating Levels for Item 15:

____5____ Tone of the report is clearly unbiased and impartial, or, negative results are presented in the text and discussed.

____4____ Tone of the report is not clearly unbiased or impartial, but the reader is able to decide, without any great difficulty, that such, in fact, is the case.

____3____ Reader has to work pretty hard to decide whether or not the tone of the report is unbiased and impartial, or, negative results are presented in the text, but are not discussed.

____2____ Tone of the report is not clearly biased and partial, but the reader is able to decide, without any great difficulty, that such, in fact, is the case.

____1____ Tone of the report is clearly biased and partial, or, negative results are presented in the tables, but are totally neglected in the text.

16. Limitations and weaknesses of the study are specified and discussed.

Scoring Points for the Five Quality of Study Rating Levels for Item 16:

__5__ All limitations and weaknesses that the reader perceives (save for certain design and statistical deficiencies which, we might assume, the researcher was unaware of) are presented and discussed.

__4__ Some limitations and weaknesses of the study are presented and discussed. The researcher may or may not neglect to mention one or more limitations or weaknesses which the reader considers fairly minor, and which the reader believes one could reasonably expect the researcher to be aware of.

__3__ Researcher mentions one or more limitations or weaknesses, but, then, in the reader's opinion, he illegitimately explains them away, or, fails to discuss them.

__2__ Researcher neglects to mention one or more limitations or weaknesses which the reader considers fairly important, and which the reader believes one could reasonably expect the researcher to be aware of.

__1__ No mention is made of any limitations or weaknesses of the study.

Rater's determination of the overall level of quality of the study

Scoring Points for the Five Quality of Study Rating Levels for the Overall Rating:

__5__ Minimal standards: If an experimental study, it had both random sampling and random assignment; if an ex post facto study, it had random sampling, (in each case, it received a 5 on item 6), and, whether experimental or ex post facto, all reliability coefficients reported were in the .80's or above (received a 5 on item 9 if scored), and, nothing beyond the reactive effects of the experimental arrangements or multiple treatment interference appeared as validity threats (received a 5 on item 11), and, the reseach problem was considered amenable to investigation, data needed to test hypotheses could be obtained through methods selected and design chosen, and methods of data-analysis selected were adequate for testing the relationships hypothesized, with no unacceptable methods being used anywhere in the analysis (received at least a 3 on each of items 1,2,7,10,and 12).

__4__ The study did not meet all minimal standards for an overall 5, but a good study, nonetheless- it received a preponderance of 4's and 5's, and few 1's or 2's, and, the latter appeared only on the less critical items (items 3,4,5, 8,and 13 through 16).

__3__ Ratings of 4 and 5 were not in the majority, or, where a majority of 4's and 5's was present, they were offset by a sizable number of 1's and 2's, or by a 2 on a critical item (though there were no 1's on the critical items, items 1,2,6,7, and 9 through 12), or, there was a sizable number of 3 ratings. Simply, in many ways, a mediocre, not really good, not really bad, study.

__2__ Something less than mediocre. Some pretty serious flaws were present, even though it didn't quite receive any 1's on any of the critical items. Some serious flaws, and not a whole lot to offset them. Or, received more than one 2 on a critical item, irregardless of how much there was to offset them.

__1__ Unacceptable methods were used to analyze the data involved in the test of the hypotheses (received a 1 on item 12), or, for experimental studies only, there was a lack of both random sampling and random assignment (received 1's on items 6 and 11), or, there were indications that the validity and/or reliability of one or more of the instruments utilized was unacceptable (received a 1 on item 9), or, there were indications that either the research problem was not amenable to investigation, the hypotheses were not testable, or the data needed to test the hypotheses could not be obtained through the methods selected or research design chosen (received a 1 on either item 1,2,7, or 10).

CAMPBELL AND STANLEY THREATS TO INTERNAL AND EXTERNAL VALIDITY

Threats to Internal Validity

The following eight extraneous variables, if not controlled in the experimental design, might produce effects confounded with the experimental effects, thus making it impossible to unambiguously attribute results observed to the factors manipulated.

1. History: those specific events which occur between successive measurements other than the experimental variables. Example: A high school English experiment, run during the month of March, involves more than one high school. One of the schools goes to the state basketball tournament. What does this do to the comparability of results from the several schools?

2. Maturation: those factors which occur naturally with the passage of time, such as growing older, growing hungrier, becoming tired, etc. Example: A Phys. Ed. teacher proposes to measure the effectiveness of a new calesthenics program by comparing the strength of his students at the beginning of the semester with their strength at the end. No control group is used.

3. Testing: often taking a pre-test will result in improved performance on post-test regardless of the intervening activity.

4. Instrumentation: changes in the procedure for collecting responses, either due to changes in the calibration of the instrument, changes occuring within observers or judges, or changes to different observers or judges. Example: One is comparing the relative amount of freedom within a classroom, first under one teaching approach and then under a second. The observer makes daily observations under the first condition, but becomes ill at the time of the switch to the second treatment. A new observer finishes the job.

5. Statistical regression: subjects whose scores are high on a pre-test will tend to have lower scores on a retest while those scoring low will tend to have higher scores on a retest.

6. Biases due to differential selection: Example: A classic English experiment (about 1930) attempted to assess the value of having every school child drink a pint of milk per day. Teachers were left with some prerogative as to who should receive the milk. A greater proportion of under-fed lower class children ended up in the experimental group with disastrous results to the interpretability of the experimental results.

7. Experimental mortality: differential loss of respondants from the comparison groups. Example: two approaches to teaching high school bookkeeping are to be compared. One approach proves much more frustrating to poorer students than the other, causing many of these poorer students to request to drop the course before the end of the experimental period. This would leave a group in one group that was, on the average, more able than the other.

---

[3] Campbell, Donald T., and Stanley, Julian C. Experimental and quasi-experimental designs for research, Chicago: Rand McNally, 1966

8. Interaction of selection with other factors: When one is forced to use groups which have not been selected by an appropriate random process, even though one may have considerable assurance that the groups may be relatively equal at the beginning of an experiment, factors which resulted in the particular composition of the groups may cause differential effects independent of the treatments. Example: a foreign language teacher decides to compare two approaches to teaching Spanish by trying one on her first period class and the other on her second period class. While no grouping is consciously practiced in assigning students to Spanish class, it is practiced in math and the best math class meets second period. Result is that all best math students who take Spanish are assigned to first period class. If first class does better, is is due to the teaching method or to a disproportionate share of able students?

Threats to External Validity

Those factors discussed by Campbell and Stanley which might jeopardize the external validity or representativeness of an experiment are:

9. Reactive or interaction effect of testing : Pre-test may cause subjects to react differently to the experimental treatments than a population of subjects who had not been pretested. Example: One is interested in trying to change students' attitudes toward minority groups. A pre-test is deemed necessary to determine if change takes place. However, having had the pre-test makes subjects more aware of the later, subtle, attempts to change attitudes than they would have been without pretest.

10. Interaction effects of selection biases and the experimental variable.

11. Reactive effects of experimental arrangements: Experimental settings may be so artificial or have so much effect as to preclude any generalization to persons exposed to the variable in non-experimental situations.

12. Multiple-treatment interference: When some subjects receive more that one treatment, prior treatments may have an effect on later treatments, since prior treatments are usually not erasable.

APPENDIX E

CONTENT CATEGORIES AND ASSIGNMENT CRITERIA

Content categories and assignment criteria

1. Locus of Control Measurement - the development of, and the presentation of reliability and validity data concerning, new locus of control scales.

2. Attribution of Causality - the attribution of responsibility or causality for actions and outcomes.

3. Academic Achievement - relationship between locus of control and academic behavior and outcomes; relationship between locus of control and achievement motivation; responses to academic success and failure; relationship between locus of control and various characteristics and features of the academic environment.

4. Social Influence Processes - responses to environmental manipulation and constraints, to include restriction of freedom of choice, and attempts to effect changes in opinions, values, attitudes, and preferences, and behaviors related to these.

5. Alcoholism and Drug Abuse - relationship between locus of control and alcohol or drug use; characteristics, behavior, and responses of those who abuse alcohol and drugs, to include responses to attempts to influence such behavior.

6. Psychological Adjustment - locus of control orientations of the psychologically maladjusted, exclusive of those individuals most appropriately described as abusers of alcohol or drugs; relationship between locus of control and factors thought to produce or to be indicators of psychological adjustment or maladjustment, to include responses to stress and other significant life events.

7. Behavior and Individual or Group Counseling - the psychotherapeutic process; responses to various psychotherapeutic approaches and techniques.

8. Self Control of Behavior - relationship between locus of control and the individual's moderation of, exercise of discipline over, his own behavior, to include the responses of the autonomic nervous system.

9. Strategy Preferences and Learning - strategies related to task performance and choice of reinforcement; the process of learning, as opposed to academic outcomes, and responses to it, to include responses to non-academic success and failure; non-academic task performance.

10. Interpersonal Perception and Processes - relationship between locus of control and Interpersonal perception; interactional processes, to include perceptions of status and or power differentials and responses to these.

11. Job Competence and Satisfaction - relationship between locus of control and work performance and satisfaction; occupational aspirations and career planning; responses to vocational counseling and vocational training.

12. Information Acquisition and Use - the acquisition, processing, and use of task or situation -relevant information and the effects of the use of such information on subsequent responses, to include task performance.

13. Cross-Cultural Comparisons - comparison of locus of control orientations across differing national groups.

14. Research on the I-E Scale - reliability, validity, and factor structure of the various locus of control scales; investigation of selected specific factors as potential threats to validity or reliability.

15. Social -Political Involvement - political participation social-political ideology and activism; reactions or responses to sociopolitical events.

16. Group Differences - comparison of locus of control orientations of differing groups within a national group, to include those defined by age, sex, race, and other personal attributes; relationship between locus of control and selected characteristics of such groups, when a comparison of two or more distinct populations is involved.

17. Personality - relationship between locus of control and other personality variables; responses of those differing across personality variables, save for those most appropriately described as psychologically maladjusted or as abusers of alcohol or drugs; relationship between locus of control and personality structure and dynamics; relationship between locus of control and self- perceptions.