SAMPLE SIZE IN TWIN STUDIES

A Thesis

Presented to

the Division of Psychology

and Special Education

EMPORIA STATE UNIVERSITY

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

By

Catherine M. Barker

August, 1988

AN ABSTRACT OF THE THESIS OF

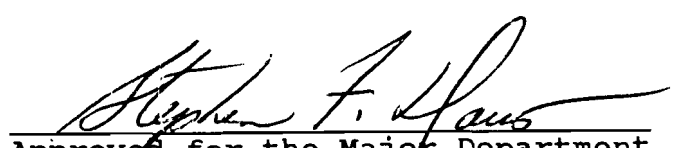
Catherine M. Barker for the Master of Science

in Psychology presented on _____

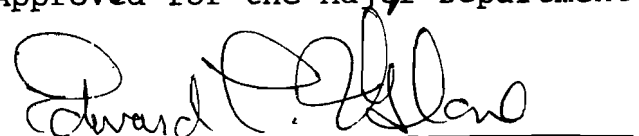
Title: Sample Size in Twin Studies

Abstract approved: Cooper B. Holmes

A study regarding sample size in human twin studies was conducted. The subjects were 70 studies collected from 6 journals that were psychologically oriented. Each issue of each journal, from the original publication date until the end of the year 1986, was examined in order to obtain twin studies. Data were then collected from each twin study as to total sample size, number of monozygotic and dizygotic twins, and number of male and female twins. In some studies it was not possible to determine sample sizes or data were not presented; thus, these studies were not included in the data. Descriptive statistics were used to examine the data for all journals combined. Results indicate that in the journals examined, and in the limited number of twin studies located, sample size in those twin studies appears to be low.



Approved for the Major Department



Approved for the Graduate Council

ACKNOWLEDGMENTS

Many individuals have helped me to reach the point in my life that made this thesis possible. The names of those to thank for their guidance, wisdom, support and love are many. To those not mentioned below; Thank you for all you have helped me to accomplish.

To Dr. Cooper B. Holmes, my advisor and thesis committee chairman, a special thank you for your guidance and assistance throughout this research.

I would also like to thank my thesis committee members. To Dr. David Dungan, for your ideas and suggestions, that helped this project take form and meaning. To Dr. Philip Wurtz for your valuable input into this project. A deep sense of appreciation also goes to Dr. Stephen F. Davis for his teaching, guiding and supportive manner that was needed throughout my stay at Emporia State University.

My warmest thanks to my colleagues and friends for allowing me into their lives. I've learned many lessons and survived many experiences while knowing you all. To Amy, Melissa, Marcia, Brian, Donna, Rosemary, Dan, Susan, Todd, and even Jennifer.

Final thanks go to my family, including Barbara L. Bartnikowski and Steven V. Barker, for their encouragement and support throughout my graduate experience. Without them, my accomplishments would take on less meaning.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Early investigation into the effects of human heredity and environment is attributed to Sir Francis Galton (Thompson & Wilde, 1973). Galton studied such characteristics as scholastic aptitude, clerical ability, and athletic skills. It was concluded by Galton that many human characteristics are strongly influenced by heredity (Thompson & Wilde, 1973; Kringlen, 1967). The role of environment was not ignored, it was just considered by Galton to have a lesser degree of influence.

Researchers have elected to continue studying the effects of heredity and environment despite Galton's strong conclusions. Through the years, researchers have refined the early methods used to assess the effects of heredity and environment. Thompson and Wilde (1973) outlined the methods of study and the characteristics studied in both animal and human research. These authors defined the major goal of these studies as being one of establishing estimates of heritability. Heritability is the variance in behavior attributed to genetic factors, when environmental factors have remained constant (Harre' & Lamb, 1983). It is further noted that the most common method used to determine these estimates is

through the use of twin studies.

There are two types of twins used in twin studies. Monozygotic (MZ) twins, commonly known as identical twins, are formed from one ovum; thus, they are genetically identical. Dizygotic (DZ) twins, commonly known as fraternal twins, are formed from two separate ova. They are no more alike genetically than are siblings (Thompson & Wilde, 1973).

Twin studies are defined as the comparison of genetically identical humans, MZ twins, to genetically similar humans, DZ twins. The purpose is to determine some degree of interaction between genetic and environmental factors in human behavior (Harre' & Lamb, 1983). The smaller variability with a particular characteristic in MZ twins is interpreted as greater genetic similarity (Harre' & Lamb, 1983). Likewise, the smaller variability with a particular characteristic in DZ twins is interpreted as greater environmental influence.

Bulmer (1970) noted that even though large amounts of twin data have been collected and reported, there has not been adequate interpretation of these data. Twin data have not been examined in a critical manner, which may lead to a discounting of the twin method or misinterpretation of the results. Bulmer stated that twin data should be interpreted with caution, as there is the possibility for several types of bias.

Smith (1974) supported the need for a closer look at the twin method of research and results. Smith pointed out that twin data and research results are not that dynamic due to the lack of consistency in twin research. This has led to a lack of comparability between twin studies. Briefly noted was the need for larger samples in order to have larger sets of twin data.

Thompson and Wilde (1973) noted that twin studies are frequently criticized for not meeting certain criteria. Much of the criticism has focused on the samples used in twin studies. Samples of twins are not representative of the total population. Twin samples are also noted for not being representative genetically and environmentally of the total population at large. Allen (1955) noted that twin samples are not representative in these two conditions due to the difficulty associated with estimating the frequency of twins. Allen suggested several methods for correcting these problems. One of these suggestions was that twins be considered as individuals and not pairs. In a later article by Allen (1965) it was specifically noted that twin samples should meet several conditions. These conditions included that the frequency of twins in the sample should be as high in proportion as the frequency of twins in the population. Age of the twins, maternal

age and race should also be considered when necessary. The twin sample should have the same proportion of same sex and opposite sex twins as are found in the population, considering age and race when necessary. Finally, the detection of twins should be uniform (Allen, 1965).

There may be other reasons for poor representation of twin samples. Martin and Martin (1975) noted that it is common for twin samples to consist of larger numbers of MZ twins and females. They attributed this in part to the greater mobility of males and the willingness of females to participate in studies. Another factor contributing to this imbalance is the fact that researchers in the past have used poor methods to determine zygosity in twin samples (Martin & Martin, 1975). This would cause inaccurate representation of MZ and DZ twins.

Lykken, Tellegen and DeRubeis (1978) referred to the problem of greater numbers of MZ twins and females as the rule of two-thirds. They noted the ratio in research is consistently that of 2:3. There are 2 females for every 3 males, and 2 MZ twin pairs for every 3 DZ twin pairs. As a solution to this Lykken, Tellegen and DeRubeis suggested that greater incentives be given to males and DZ twin pairs to participate in twin studies.

A further problem in obtaining a representative sample is that of selection effects. Martin and Wilson

(1982) contend that those less able to participate are often excluded from a study. This type of sample is named the truncated sample. Frequently researchers obtain their samples by non-random selection, thus excluding some subjects. When this truncation is ignored, correlation coefficients are lowered (Martin & Wilson, 1982). This leads to a bias in the estimation of proportions of genetic and environmental variance. Studies in educational achievement were cited by Martin and Wilson supporting their contention that samples are not randomly selected, thus effecting results.

Torgersen (1983) conducted a study of the influence of sampling variation upon the difference in concordance rates between MZ and DZ twin pairs. He noted that past research of neurosis utilized samples of twins that were highly selected, those samples being obtained mostly from hospitals. It is suggested that the best sample would be one utilizing a nation-wide population (Torgersen, 1983). The high infant mortality rate in twins is noted as being a problem in all twin research and no suggestions are made to correct for this problem.

Fuller and Thompson (1960) have suggested that in order to collect a representative sample of twins it is best to use the total registry of twin births. This would insure a homogeneous sample that is large enough. They did not suggest any particular size for samples.

Fuller and Thompson did suggest that twin pairs should be selected from the same population by age, culture and geographical location. Subjects should only be dropped from a study due to unavailability or accident to one member of a twin pair.

There are, however, some problems in using twin registries. Kendler and Holm (1985) contended that there is a selection bias in the methods used by twin registries that reflects higher rates of major diseases in DZ twins. The registry requires that both members of a twin pair pass a medical examination in order to be included. Kendler and Holm concluded that this differential enrollment in twin registries will significantly effect results of twin studies and should not be ignored.

As a result of the various difficulties in obtaining representative twin samples, sample size has been affected. Falconer (1960) demonstrated the importance of adequate sample size in twin studies. He noted that when considering gene frequencies, large populations are stable. Without such factors as migration, mutation and selection, genes and genotype frequencies are consistent across generations. The same is not true with small populations where the gene frequencies are subject to fluctuations. Given this, if the number of subjects taken from a population is too

small, there will be genetic differences in that sample from that population (Falconer, 1960). If the sampling is not large and random, an accurate sample will not be drawn.

Several other researchers have remarked on inadequate sample size in twin studies (Christiansen, 1973; Cochrane & Duffy, 1974; Dworkin, Burke, Maher & Gottesman, 1981; Kringlen, 1986; Parker, 1986). Klein, DeFries and Finkbeiner (1973) noted that there has been a failure in the past to consider sample size which has led to inadequate reporting of results. Specifically, heritability rates have been reported without standard errors being given. Klein, DeFries and Finkbeiner (1973) provided some statistical methods that can be used to help accurately estimate heritability rates. They concluded that a sample of 400 families, with 2 offspring per family, is recommended as the minimum sample size when studying heritability.

Carey, Goldsmith, Tellegen and Gottesman (1978) concluded that when twin studies do not have a large enough sample size the standard error for accurate estimates of genetic parameters will be reduced. Carey et al. (1978) suggested that future research should look towards increasing numbers of DZ twins in samples for reasons cited earlier (see Lykken, Tellegen & DeRubeis, 1978; Martin and Martin, 1975). When large samples can

be obtained and data from independent studies are available, inferences of differences between MZ and DZ correlations may be explored for genetic factors (Carey et al., 1978).

McGue and Bouchard (1984) noted that inadequate sample size in twin research may result in under or over estimation of correlations, depending upon the statistical analysis used regarding age-sex variables. Sample sizes examined by McGue and Bouchard (1984) ranged from 25 to 200. Results suggested that even when small sample sizes (25 or less) are used, corrections can be made as they relate to age and sex.

Eaves (1972) noted that classical twin studies are based on sample sizes of less than 100 twin pairs. This appears to be adequate when examining within family genetical and environmental influences. Larger samples are needed for the purpose of gaining more information on the genetical and environmental interaction (Eaves, 1972). Larger samples are also needed for other types of twin-related research methods. Eaves concluded that earlier computer simulated studies, as well as his own study, indicated that much larger sample sizes are needed in order to extend our current knowledge.

Cattell, Schuberger, Ahern and Kameoka (1981) supported Eaves contention that larger sample sizes are needed. With larger samples it is suggested that specific

theories can be tested (Cattell et al., 1981). In psychological research it is often the case that one body supports a conflicting view (Holmes, 1979). Sample size may be a contributing factor in producing these conflicting results. McNemar (1940) discussed what an appropriate sample size is for any type of research and concluded that there is no particular number that can be recommended.

Hay (1953) also noted that many times in research small sample size is a problem. Small samples can be adjusted statistically, yet this does not take into account that there may be characteristics associated with small samples that aren't present in larger samples. Thus, no amount of playing with the statistics will benefit a small sample.

Stolurow and Frincke (1966) examined the use of small samples in developing programmed learning materials. They concluded that it is best to use as large a sample as is practical in order to increase the statistical power of a study and to decrease errors.

Holmes (1979) has noted that sample size in psychological research is an issue that calls for closer examination. In order to gain further knowledge regarding sample size, Holmes examined selected APA journals for the complete years of 1977 and 1955. It was concluded by Holmes that sample size did not appear to be large. Further research by Holmes, Holmes and Fanning (1981)

supported the earlier results that sample size is not large in non-APA journals for the year 1977.

While sample size is an elusive issue, it has been shown to be a problem when trying to draw conclusions from twin studies. However, there are no specific studies dealing exclusively with sample size in twin research. It is the purpose of this paper to survey twin studies, in selected journals, in order to provide data regarding the typical sample size used in twin studies.

CHAPTER 2

METHOD

Subjects

The subjects for the study consisted of 70 human twin studies, located in 6 journals. The journals, years surveyed, and number of studies collected from each are as follows: The American Journal of Human Genetics, 1949-1986, 27 studies; American Journal of Psychiatry, 1921-1986, 14 studies; Behavior Genetics, 1970-1986, 24 studies; Developmental Psychology, 1969-1986, 4 studies; Human Development (Vita Humana, 1958-1964), 1958-1986, none; Journal of Personality and Social Psychology, 1965-1986, 2 studies. All journals were surveyed from the original publication date until the end of the year 1986. For example, the American Journal of Human Genetics has been in print since 1949 to the present date, and was surveyed from 1949 to 1986. At the time of data collection, 1986 was the last complete year.

These journals were selected based upon their being frequently cited by authors of twin research in journals and text books. The Directory of Publishing Opportunitites in Journals and Periodicals was also used to confirm the areas of research accepted by these journals.

There are two ns for this study. The first n is

the total sample size, or total number of subjects analyzed, regardless of the number of groups used in each study. The second n represents the individual samples or the size of each group that makes up the total. For example, if a study consists of a total of 180 subjects, with 3 groups of 60 subjects in each group, the total n is 180 and the individual sample size n is 60 (three times). This approach is utilized as conclusions from results are often based upon subjects classified under specific conditions and not upon the sample as a whole.

Procedure

Each issue of each journal, during the dates of publication previously noted, was examined in order to obtain twin studies. Each twin study, by definition those comparing MZ and DZ twins, was recorded. Data recorded for each study included the total sample size studied, individual sample size (the number of MZ, DZ and twins of unknown zygosity), and the number of males and/or females studied. In some studies it was not possible to determine sample size, or data were not reported. These studies were not included in the data.

CHAPTER 3

RESULTS

The results are presented as descriptive statistics. The mean, median, mode, standard deviation, 25th percentile, 75th percentile, and range of sample sizes are reported for all journals combined, i.e., there is no journal by journal comparison. The data are reported separately for total n and individual ns. The data in Table 1 and Table 2 refer to twin pairs. The data in Table 3 refer to twins as individuals.

Table 1 presents the results of the total sample size of twin studies examined. The total sample n is 70 studies. The remaining figures present the number of pairs of twins respective to each statistical formula specified. For example, the mean number of twin pairs is 652 pairs. Table 1 also presents statistics calculated without extreme scores. That is, studies with 500 or more twin pairs were excluded, leaving a total of 64 studies. The second column in Table 1 presents the recalculated figures excluding those studies with 500 or more twin pairs.

Table 2 presents the results for the individual sample n taken from the 70 twin studies. The individual n is composed of MZ and DZ twin pairs and twin pairs of unknown zygosity.

Table 3 presents the results of the total sample of 70 twin studies when classified by gender. Subjects were

Table 1

Total Sample Size of Twin Studies Collected from Six Journal Publications

n	70	64[a]
Mean	652	134
SD	2,581	123
Median	102	87
25th Percentile	46	46
75th Percentile	262	191
Mode	46	46
Range	15,905	478
	(1,5909-4)	(482-4)

Note. All statistics refer to pairs of twins.

[a]This column calculated after excluding studies with $n \geq 500$ pairs.

Table 2

Individual Sample Size in Twin Studies by Zygosity

n	151[a]	142[b]	140[c]	133[d]
Mean	302	73	269	71
SD	1,060	77	1,051	71
Median	54	48	51	47
25th Percentile	21	20	21	20
75th Percentile	121	96	117	90
Mode	13	13	13	13
Range	7,224 (7,225-1)	460 (461-1)	7,224	460

Note. All statistics refer to pairs of twins.
[a]This column represents all twin pairs in all studies.
[b]This column represents all twin pairs in studies containing $\geq$ 500 pairs. [c]This column represents twin pairs excluding twins of unknown zygosity. [d]This column represents twin pairs excluding twins of unknown zygosity and excluding studies containing $\geq$ 500 pairs.

ble 3

Total Sample Size in Twin Studies by Gender

	Male		Female	
n	44	36[a]	33	29[b]
Mean	930	117	306	108
SD	4,449	95	829	83
Median	100	89	94	88
25th Percentile	58	43	41	35
75th Percentile	282	146	188	155
Mode	92	92	27,49, 50,94,120, 134,213	27,49, 50,94,120 134,213
Range	29,952 (29,962-10)	538 (548-10)	4854 (4872-18)	314 (332-18)

Note. All statistics refer to individuals.

[a]This column represents individual samples of male twins excluding studies containing ≥ 500 individuals. [b]This column represents individual samples of female twins excluding studies containing ≥ 500 individuals.

were classified as male, female, or unreported gender. It was not possible to determine the number of males and females in three studies, though in all three data for males and females are reported by the authors. As a result, data from these three studies were excluded. Of all the studies examined for this research project, gender was clearly identified in 44 individual samples of male twin pairs. Gender was clearly identified in 33 individual samples of female twin pairs.

CHAPTER 4

DISCUSSION

The results from this research project present data that reflects the sample size used by researchers studying twin populations along a variety of characteristics. Specific journals were examined and a limited number of twin studies were located. No conclusions are made regarding the size of twin samples as there are no set figures required for a study to be statistically sound. The examination of sample size in research, much less twin research, has not been extensively studied. A variety of descriptive statistics were used here in order to provide the most accurate evaluation of the data. The mean is subject to fluctuations caused by extreme scores. Thus, the median and percentiles were also calculated. For further information, the mode and range were provided. It is hoped that these will provide a variety of information to be used and compared in future studies of twin samples. In order to aid in further research it is important to note a few observations.

Data were not always clearly identifiable. It was often necessary to refer to tables in order to clarify data. This was not always successful in determining sample size and led to the exclusion of data. Several studies referred the reader to specific samples used previously, when the data had been published elsewhere, and were not

readily available. These potential data were also excluded.

It is interesting to note that researchers did not always clearly indicate the zygosity or gender of twin samples. In most studies these data were determined by closer examination of text and tables. In some studies researchers collected samples of twin pairs in which the zygosity of a minority of twin pairs were not reported or could not be determined by methods used at the time. Researchers included these data in their results, making their results questionable. The same is true of the reporting of data regarding the gender of twin pairs.

REFERENCES

References

Allen, G. (1955). Comments on the analysis of twin samples. Acta Geneticae Medicae at Gemellologiae, 4, 143-159.

Allen, G. (1965). Twin research: Problems and prospects. In A. G. Steinberg, & A. G. Bearn (Eds.), Progress in medical genetics. (pp. 242-269). New York: Grune and Stratton.

Bulmer, M. G. (1970). The biology of twinning in man. London, England: Oxford University Press.

Carey, G., Goldsmith, H. H., Tellegen, A., & Gottesman, I. I.(1978). Genetics and personality inventories: The limits of replication with twin data. Behavior Genetics, 8(4), 299-312.

Cattell, R. B., Schuberger, J. M., Ahern, F. M., & Kameoka, V. (1981). The heritability of fluid and crystallized intelligences: By the MAVA design and OSES analysis. Australian Journal of Psychology, 33(3), 355-374.

Christiansen, K. (1973). Mobility and crime among twins. International Journal of Criminology and Penology, 1, 31-45.

Cochrane, R., & Duffy, J. (1974). Psychology and scientific method. Bulletin of the British Psychological Society, 27, 117-121.

Dworkin, R. H., Burke, B. W., Maher, B. A., & Gottesman, I. I. (1976). A longitudinal study of the genetics

of personality. Journal of Personality and Social Psychology, 34(3), 510-518.

Eaves, L. J. (1972). Computer simulation of sample size and experimental design in human psychogenetics. Psychological Bulletin, 77(2), 144-152.

Falconer, D. S. (1960). Introducation to quantitative genetics, New York: Ronald Press.

Fuller, J. L., & Thompson, W. R. (1960). Behavior genetics. New York: Wiley.

Harre', R., & Lamb, R. (Eds.). (1983). The encyclopedic dictionary of psychology. Great Britian: Basil Blackwell.

Hay, E. N. (1953). A note on small samples. Journal of Applied Psychology, 37, 445.

Holmes, C. B. (1979). Sample size in psychological research. Perceptual and Motor Skills, 49, 283-288.

Holmes, C. B., Holmes, J. R., & Fanning, J. J. (1981). Sample size in non-APA journals. The Journal of Psychology, 108, 263-266.

Kendler, K. S., Holm, N. V. (1985). Differential enrollment in twin registries: Its effect on prevalence and concordance rates and estimates of genetic parameters. Acta Geneticae Medicae et Gemellologiae, 34 (3-4), 125-140.

Klein, T. W., DeFries, J. C., & Finkbeiner, C. T. (1973). Heritability and genetic correlation: Standard errors of estimates and sample size. Behavior Genetics, 3(4),

355-364.

Kringlen, E. (1967). Heredity and environment in the funcitonal psychoses. London, England: William Heinemann Medical Books.

Kringlen, E. (1986). Status of twin research in functional psychosis. Psychopathology, 19, 85-92.

Lykken, D. T., Tellegen, A., & DeRubeis, R. (1978). Volunteer bias in twin research: The rule of two-thirds. Social Biology, 25(1), 1-9.

Marquis Academic Media. (1981). Directory of publishing opportunities in journals and periodicals. (5th ed.). Chicago, IL: Marquis Academic Media.

Martin, N. G., & Martin, P. G. (1975). The inheritance of scholastic abilities in a sample of twins: I. Ascertainment of the sample and diagnosis of zygosity. Annals of Human Genetics,39, 213-218.

Martin, N. G., & Wilson, S. R. (1982). Bias in the estimation of heritability from truncated samples of twins. Behavior Genetics, 12(4), 467-472.

McGue, M., & Bouchard, T. J. Jr. (1984). Adjustment of twin data for the effects of age and sex. Behavior Genetics, 14(4), 325-343.

McNemar, Q. (1940). Sampling in psychological research. Psychological Bulletin, 37, 331-365.

Parker, G. (1986). Validating an experimental measure of parental style: The use of a twin sample. Acta

Psychiatrica Scandinavica, 73, 22-27.

Smith, C. (1974). Concordance in twins: Methods and interpretation. American Journal of Human Genetics, 26, 454-466.

Stolurow, L. M., & Frincke, G. (1966). A study of sample size in making decisions about instructional materials. Educational and Psychological Measurement, 26, 643-659.

Thompson, W. R., & Wilde, G. J. (1973). Behavior Genetics. In B. B. Wolman (Ed.), Handbook of general psychology. (pp. 206-229). Englewood Cliffs, NJ: Prentice-Hall.

Torgersen, S. (1983). Genetics of neurosis: The effects of sampling variation upon the twin concordance ratio. British Journal of Psychiatry, 142, 126-132.