The Cochran-Mosteller-Tukey Evaluations of the Kinsey Report Revisited – FIU-01/08/20013

Ive Barreiros

Abstract: This article revisits the assessment that a committee specially appointed by the American Association of Statistics, made of Alfred Kinsey's book "The Sexual behavior of the Human Male", published in 1948.

The committee, formed by William G. Cochran, Frederick Mosteller, and John W. Tukey, critically analyzed the methodology of the research while maintaining a somewhat benevolent attitude, justifying mistakes of the researcher given the circumstances and the nature of the investigation.

In spite of objections of a statistical nature, the "Kinsey Report" had great influence on the education of adolescents and children, laws on sexual offences and even the private behavior of many Americans. Over time, many analysts have been discovering more and more details on Kinsey's modus operandi, and the influence his own personality exercised into his work. The article analyzes the original assessment and raises questions about whether the Statistical Committee should have clarified, while using statistical means, a more open and clear set of serious objections, both technical and ethical, to the "Kinsey Report."

Contents

"If his science was flawed — or worse yet, an outright deception — then our culture's attitudes about sex is not just wrong morally but scientifically as well." (From Kinsey's Secret: The Phony Science of the Sexual Revolution, by Sue Ellin Browder)

I. Introduction

Professor Oscar Varsavsky, at the School of Exact Sciences of the University of Buenos Aires, addressing students of mathematics close to graduation used to say: "Do not forget!...most of you will be supporting a physicist, an economist, a sociologist,... do not forget that you will be providing formality, scientific rigor, to the work of others." This statement contained a double lesson: the "supporting" nature of our role and the call for providing scientific seriousness to research. He also used to say: "Work in a detailed, meticulous and strict way, even if they would not pay you."

That applies, I think, to statisticians. Their function is providing support and scientific rigor … with a twist. In fact, rigor in statistics can be reached only by facing a battle on a three folded front: sampling, measuring and the final inference; all mutually related.

When working as a statistician in the applied field, you soon learn that the X variable of the inference theorems plays "hard to get" most of the times. You have to face the reality of defective frames, reluctant subjects, non -response, incomplete responses, refusals, uncertain expansion coefficients, and the like. (If my professors ever told me that non sampling errors are more frequent and less measurable than the sampling ones, I probably did not pay attention or chose not to listen).

You also must become aware of the delicate position of the statistician working as consultant, survey statistician, statistical designer, and evaluator. By supporting scientific rigor to processes of obtaining and analyzing data they provide information and support to many decisions of individuals and society.

The case I present in this paper is one that has puzzled me for years. It is the evaluation carried out by William Cochran, Frederick Mosteller and John W. Tukey, of *Sexual Behavior in the Human Male*, the first of so called "Kinsey reports." I think that this evaluation illustrates a situation in which well-intended statisticians were not clear enough about their role of claiming scientific rigor in research and informing society of the real value of a specific work.

In 1948 and 1953, the United States was shocked by events that some observers compared to the explosion of the atomic bomb: the publication of *Sexual Behavior in the Human Male and Sexual Behavior* in the Human Female, respectively, popularly known as "the Kinsey Reports."

Extensive data on sexual acts, compiled by the Indiana University zoologist Alfred Kinsey and a team of researchers, graphically presented the results of interviews with thousands of American men and women. By his shocking findings, Kinsey's work seemed to expose that much of Americans' sexual activity took place outside of marriage, and that the majority of the nation's citizens had violated accepted moral standards as well as state and federal laws in their pursuit of sexual enjoyment. (Duke, 2008)

Kinsey's work is one of the most famous and most ballyhooed research projects of the 20th century. Establishing that fraud took place in his research has not been particularly difficult, as I will soon show. However, this major piece of deception was passed off as valid and innovative science for more than fifty years.

In fact, the Kinsey's ideas permeated most of the scientific and policy literature of the postwar years. Amazingly often the findings of the reports, along with public and media responses to them, have been featured in discussions of American society and national identity. (Reumann, 2008)

The Committee for Research in Sexual of the National Research Council requested from the American Statistical Association (ASA) an evaluation of the first Kinsey's report.[1] The report by the American Statistical Association covers only the Kinsey's book about males. Its title is *Statistical Problems of the Kinsey Report*, and was published by ASA in 1953.

The ASA report pointed out serious statistical deficiencies regarding sampling techniques, non-sampling errors, statistical methods, interpretation, and presentation of results. However, some segments of the scientific community and the public in general were left with the impression that the findings were scientific and the methods were either the best or the only possible actions available, given the delicate topic investigated and the resources of the time.

It is not the intention here to analyze the cultural and social circumstances and dilemmas that are behind the massive acceptance and influence of the Kinsey reports in psychology, education and public policy during the last fifty years.[2] This article is about the ASA evaluation report and the failure

[1] This committee provided the major funding for the first Kinsey report.

[2] Those interested in details about the impacts of the Kinsey report will find interesting the book American Sexual Character by Miriam G. Reumann. University of California Press. 2008. Those interested in details about the influence of the Kinsey reports in law, education and public opinion are recommended to read the article: "Restoring Legal Protection for Women and Children", by Linda Jeffrey. The State Factor. American Legislative Exchange Council. 2004

of its authors, all already prestigious statisticians, to clearly communicate to the scientific community and the general public the real magnitude and nature of Kinsey's errors.

I have tried to present and show the overwhelming evidence that ASA evaluators had in their hands. My conclusion is that, by restricting the scope of their work and keeping an unjustifiable and condescending attitude toward Kinsey and the sex research, they missed a precious opportunity to:

- fully identify and asses the purpose of the study,
- properly secure the protection of the subjects involved, and
- openly object to the validity of the findings and their application to public policy

That failure occurred in spite of the fact that the evaluation team had all the necessary elements to categorize the first Kinsey report as a complete methodological fraud and made that truth available to everybody. They were clearly stated in more than 100 identifiable and specific critiques written by a team of distinguished research specialists of recognized capacity and professional prestige. The present work is a classification of those critiques according to the statistical topic whose requirements are disregarded or plainly violated.

II. Background

The book *Sexual Behavior in the Human Male*[3] (KPM), by Alfred C. Kinsey, Wardell B. Pomeroy and Clyde E. Martin[4], was published in 1948. Major funding of Kinsey's work was provided by the Committee for Research in of Sex (CRPS), pertaining to the National Research Council.

Kinsey had earned a PhD at Harvard and became a biology professor at Indiana University where he wrote biology textbooks and a book about gall wasps. He was an entomologist by training; a foremost authority on gall wasps. It was at Indiana University that Kinsey's interest in sex research arose after he was asked to participate in a sex education course. This course was to prepare students for fulfilling marriages.

Kinsey's liberal attitudes resulted in his being quickly replaced by the university administration in teaching the sex education class. Yet Kinsey's interest in sex research grew and he began the research that eventually led to the formation of the Institute for Sex Research at Indiana University.

[3] Kinsey AC, Pomeroy WB, Martin C.E. *Sexual Behaviour in the Human Male*. Philadelphia: WB Saunders,1948.

[4] Referred here as KPM

It was through this institute that in 1948 he published the book, *Sexual Behavior in the Human Male.* This was followed five years later *by Sexual Behavior in the Human Female*.[5]

5300 white males and 5940 white females provided almost all the data, with the majority of participants being younger white adults with some college education. (This part of the sample is referred to as the "College Sample.") Kinsey tried to compensate for volunteer bias in his sample by interviewing 100% of the individuals available in a given organization or group. Approximately 25% of the sex histories came from these "100% groups." Kinsey did not believe a random sample was possible. He used in-depth, face-to-face interviews by highly trained interviewers. In each history a subject would be questioned on up to 521 items, depending on his/her specific experience (the average in each case being near 300). Histories covered social and economic data, physical and physiologic data, marital histories, sexual outlets, heterosexual histories, and homosexual histories.[6]

Once published, the KPM report generated "some questioning" in published articles.[7] Thus the Committee for Research in Problems of sex, of the National Research Council, asked the American Statistical Association to provide council, through evaluation and advice, regarding research methods to the KPM work. The request also attempted to assure unquestioned acceptance for the second volume, then in preparation.

The complete report, entitled *Statistical Problems of the Kinsey Report on Sexual Behavior in the Human Male*, was published in its entirety by the American Statistical Association in 1954. The authors, William Cochran, Frederic Mosteller, and John W. Tukey were appointed by the Association's Commission on Statistical Standards. They had the assistance of W.O. Jenkins. It will be referenced from now on as CMT.

The committee had the cooperation of Kinsey with the inclusion of some visits to the Institute of Sex Research, Inc. at the University of Indiana. Also the authors went through the interviewing process that Kinsey used in gathering the data for his book. The entire work was really the result of a symposium including papers by Kinsey, himself, and discussion of criticisms by 9 other statisticians.[8]

---

[5] Kinsey AC, Pomeroy WB, Martin CE, Gebhard PH. *Sexual Behaviour in the Human Female*. Philadelphia: WB Saunders, 1953.

[6] The Kinsey Institute official information

[7] Statistical Problems on the Kinsey Report. pag 5

[8] In a letter to Wilks, President of ASA, dated August 28 1950, Kinsey expresses his appreciation for the willingness of ASA to undertake the evaluation, and offers total cooperation. However, he points out that the resulting recommendations will be taken into account in future works, but will affect the forthcoming report (on female sexuality) only with regard to limitation of the methods and conclusions.

Kinsey's Shocking Findings

The findings of the KPM report that motivated ASA to ask for council and evaluation from ASA were related to:

- A high level of sexual activity
- A small change from older to younger generations
- A strong relation between activity and socio-economic class
- A relation between activity and changes of socio-economic class

After a long period of assessment, involving many meetings with Kinsey and his team, a detailed report by the review group of three -Cochran, Tukey and Mosteller (CMT) - was published. The CMT report contains basically an evaluation by the main authors, and seven appendices. Appendices A, B, and C are specific to KPM and the rest are methodological considerations.

A. More than 120 comments by six designated "reviewers". (See Summary in Appendix 2)
B. A comprising of KPM with 6 other sex researches (See Summary in Appendix 3). Five of the commentators were suggested by Kinsey himself.
C. Proposed further work
D. Considerations on sampling (theoretical)
E. The Interview as seen by CMT
F. Consideration about accuracy
G. Principles of Sampling

The Statisticians' Benevolent Evaluation

The final conclusion of CMT report was:[9]

"*All of these KPM set well as established conclusions. All are subject to unknown allowances for:*

- *difference between reported and actual behavior*
- *non probability sampling*
- *involving volunteering"*

It seems an extremely diplomatic way of saying that the KPM deserved a clear and firm negative evaluation from the highest statistical authorities in the nation. They were really saying that Kinsey's findings were subject to both sampling and non-sampling errors of unknown magnitude, although that was a too technical way of expressing the seriousness of the flaws. The evaluation failed to

[9] CMT, page 39

clearly express the magnitude and seriousness of those errors and the vagueness of the diagnosis left many to believe whatever they could or wanted.

Somehow the massive amount of evidence of errors found by the nine specialists was not clearly transmitted by CTM to the general public and the evaluation failed to warn the public against a pseudo-scientific work plagued with methodological as well as ethical objections.

America Gets Crazy about Dr. Kinsey

When Kinsey and his coworkers published *Sexual Behavior in the Human Male* in 1948 and *Sexual Behavior in the Human Female* in 1953, they turned middle-class values upside down.

The Kinsey Reports had a shocking impact on the American public. Almost overnight, Sexual Behavior in the Human Male converted the obscure mid-western college professor into a national celebrity, and his name became a code word for all things sexual. He quickly earned a notorious place in popular American culture. (Browder, 2008)

Many traditionally forbidden sexual practices were, according to Kinsey, surprisingly commonplace:

- 85 percent of men and 48 percent of women said they'd had premarital sex,
- 50 percent of men and 40 percent of women had been unfaithful after marriage,
- 71 percent of women claimed their affair hadn't hurt their marriage, and a few even said it had helped,
- 69 percent of men had been with prostitutes,
- 10 percent had been homosexual for at least three years, and
- 17 percent of farm boys had experienced sex with animals.

Implicit in Kinsey's report was the notion that these behaviors were biologically "normal" and hurt no one. Therefore, people should act on their impulses with no inhibition or guilt.[10] The 1948 report on men came out to rave reviews and sold an astonishing 200,000 copies in two months. Kinsey's name was everywhere from the titles of pop songs ("Ooh, Dr. Kinsey") to the pages of *Life, Time, Newsweek*, and the *New Yorker* (Browder, 2008).

Albert Deutsch, writing in *Harper's* magazine (December 1947, p. 494) said that it "explodes traditional concepts of what is normal and abnormal, natural and unnatural in sex behavior." *Look*

[10] Speaking at a 1955 conference sponsored by Planned Parenthood, Kinsey claimed that of all pregnant women, roughly 95 percent of singles and 25 percent of those who were married secretly aborted their babies. He gave scientific authority to the notion that abortion was already a common medical procedure — and should thus be legal.

magazine (December 9, 1947, p. 106) said the Kinsey team had produced a "social atomic bomb" that "may have a tremendous effect on the future social history of mankind". (Gordon and Court, 1992). The media effectively sold Dr. Kinsey as "presenting facts". He was compared to Darwin, Galileo, and Freud (Browder, 2008).

In a special 1990 issue of *Life* magazine, Kinsey was named as one of the 100 most important Americans of the 20th century and his conclusions were described as reflecting main stream American sexuality (Gordon and Court, 1992).[11]

Kinsey is often called the "father of the sexual revolution," some kind of modern prophet who redefined the sexual mores of everyday Americans. The science that launched the sexual revolution has been used for the past 50 years to sway court decisions, pass legislation, introduce sex education into the school system, and even push for a redefinition of marriage. Kinseyism was the very foundation of this effort (Jeffery, 2004).

Time Did Not Serve Kinsey Well

In spite of the mass media and popular approbation, there were many scientists who perceived the conclusions as flawed. Among them, the most prominent are anthropologists Margaret Mead and Ruth Benedict; Stanford University psychologist Lewis M. Terman; Karl Menninger, M.D. (founder of the famed Menninger Institute); psychiatrists Eric Fromm and Lawrence Kubie; cultural critic Lionel Trilling of Columbia University, and countless others. (Browder, 2008)

It was believed at the time Kinsey was a scrupulous and disinterested scientist during sex research. Time and study of Kinsey and of the Institute for Sex Research has shown otherwise. Besides looking critically at his research and how it was conducted, there were questions about Kinsey's own sexuality and sexual life. (Houston, 2007)

Gradually, the professional environment was becoming aware of the Kinsey fraud. "It has long been recognized that one of the greatest faults of the Kinsey research was the way in which the cases were selected: the sample is not representative of the entire U.S. population or any definable group in the population. This fault limits the comparability and appropriateness of the Kinsey data as a basic for calculating the prevalence of any form of sexual conduct." (Turner, Miller, and Moses, 1989)

---

[11] Kinsey and his team always appeared as typical middle-class Americans in publicity photographs, wearing suits and ties and posing with their wives and children whenever possible. Parading the book under the respectable cover of science, coupled with Rockefeller-connected mass media affiliations, the unconventional research of the so-called "All American", Kinsey team seemed acceptable, even state-of-the-art. (Brinkman, 2005)

Many analysts in the last 50 years have claimed that the research was not only fraudulent in its methodology, but objectionable in its practices and even unlawful in its procedures. There have been claims that the whole work was ideological motivated and took research interest and confidentiality beyond the rule of law.[12] "Kinsey was not merely presenting data in his first Report - he was making a point, a point he himself was clear about long before he handed out his first questionnaire. This colors things." (Archer, 2002)

"No one knew at the time, of course, Alfred Kinsey's impetus for embarking on his monumental and epoch-shifting study of human sexuality came from a desire to justify his own sexual thoughts and practices." (Archer, 2002)

It would be an unnecessary and very cumbersome task here to go over the many accusations of Kinsey's depravity, scientific fraud and negative impact his reports had upon society.[13] They have been exposed in many publications, a congress investigation and two movies (Gordon and Court, 1992). An interesting testimony about the influence of Kinsey's mentality in counseling services on college campuses is provided by an anonymous author in the book *Unprotected*.[14]

On the other hand, it is not that there was some truth contained in Kinsey reports. It is impossible to deny that sexuality, as an important component of personhood, is certainly a subject worthy of scientific study.

It is also credible that many people preach sexual purity while secretly behaving altogether differently in their private lives. Nevertheless, individuals and society need to make decisions based upon solid scientific research, besides their own conception of the human person and interest. But that was not the kind of research work that Kinsey and his team carried out.

---

[12] Fraud of the Century? .Gordon Muir and John H. Court. May 1992 Edition Of The Catholic Medical Quarterly.

[13] A comprehensive expose of Kinsey is presented in the book by Judith Reisman *Kinsey: Crimes & Consequences*.

[14] "Unprotected", by Anonymous, MD. Penguin Group. New York: 2006.

III. The Evaluation by Others

The Reviews by Field Experts

The final evaluation by Cochran, Mosteller and Tukey (CMT) had to be based in the conclusions contained in six "reviews" done by some recognized authorities of the time. The six reviews of the KPM report selected by CMT for detailed discussion are:

- "Statistics of the Kinsey Report," Wallis W. A. The journal of the American Statistical Association, Vol. 44 (1949).
- "Sexual Behavior of the American Male: A special review of the Kinsey report," Goldstein, J. and Pastore, N., The Journal of Psychology. Vol. 26 (1948).
- "An appraisal of Some Methodological Aspects of the Kinsey Report," Wallin, P., American Sociological Review, Vol. XIV, No.2 (1949).
- "Kinsey's Sexual Behavior of the Human Male"; some comments and criticisms." Terman, L.M., Psychological Bulletin, Vol. 45, (1948).
- "An Evaluation of 'Sexual Behavior of the Human Male'," Hobbs, A. H., and Lambert, R. D. American Journal of Psychiatry, Vol. 104 (1948).
- "The Kinsey Report and Survey Methodology," Hyman, Herbert H., and Sheatsley, Paul B. International Journal of Opinion and Attitude Research, Vol. 2 (1948).

Kinsey himself suggested the first five, while the sixth, by Hyman and Sheatsley, was added because the authors' recognized experience in opinion polls.

It is a known fact among contemporary critics of Kinsey, that the sample of individuals he used in his study lacked the minimal requirements to allow an inference on the behavior of "white American man," as he claimed. In reality, none of his evaluators who scrutinized his work concluded a clear structure of the Kinsey' sample, such as number of subjects, number of histories, and exact account for his so called "100% groups," (that the evaluators generously called "clusters.")

However they had a clear impression that considerable number of histories came from inmates, mental health patients, male prostitutes and other individuals who probably gave their histories in very specific circumstances (which should not be confused with volunteering).[15]

[15] Wallis (CTM, pg. 48) and Wallin (CTM, pg. 49) refer to the problem of understanding the composition of the sample. Terman (pg. 49), clearly states that there is not information to judge representativeness regarding volunteering and definition of the 100% groups.

The type of critics the evaluators pointed out in the CTM goes far beyond the representation problem and they definitely put in evidence other aspects of the Kinsey's style as a researcher. Appendix 2 presents a summarized version of the critical comments. They are classified[16] according to specific aspects of the evaluation. Those aspects correspond to classical aspects of research to be considered worth of statistical value:

- Sampling
- Interview
- Report
- Data presentation
- Interpretation
- Stability
- Measurement
- Presentation of data
- Interpretation

As shown in Chart #1, out of 123 comments from the reviewers, 74 (or 59%) refer to data presentation, the measurement, and data interpretation. Out of the remaining 41%, the two topics receiving more criticisms are the sampling problem and the interview.

Chart 1. Commentator and Review's Field

Topic	Goldstein and Pastore (Psychology)	Hobbes and Lambert (Psychiatry)	Hyman and Seatsley (Surveys)	Terman (Psychology)	Wallin (Sociology)	Wallis (Statistics)	**Total**
Sampling	1	0	2	4	5	3	**15**
Interview	1	1	2	5	3	2	**14**
Report	0	0	0	1	4	0	**5**
Stability	0	0	1	1	1	3	**6**
Measurement	5	0	1	3	12	5	**26**
Statistical Techniques	1	0	0	1	0	7	**9**
Data Presentation	9	3	0	1	1	14	**28**
Interpretation	4	4	0	8	1	3	**20**
Total	**21**	**8**	**6**	**24**	**27**	**37**	**123**

Source: *Statistical Problems of the Kinsey Report*. Appendix A

[16] This classification was done by Mr. Kimball Romney for the CMT report.

Regarding reviewers, 109 critiques (89%) were pointed out by Wallis, Wallin, Terman and Goldstein and Pastore. Wallis, the reviewer with a maximum number of comments (30%), and Goldstein-Pastore (17%) seems to be particularly concerned about data presentation. Wallin pointed out measurement especially, while Terman (20%) paid more attention to interpretation.

Notice that in Data Presentation (Appendix 2) "data" refers to information about every single aspect of the research. It is apparent that the reviewers had a very hard time trying to understand the details of the size and structure of the sample, the wording of questions, the content of the interview, the expansion of sample to population, and the final results.[17]

The comments under the topic called Measurement refer mainly to deficiencies in the tests for validity and reliability, insufficient information when checks are done, the use of smoothing trends, etc. It seems that the commentators recognize genuine attempts by KPM to validate their measurements, but find them poorly documented and not comprehensive. The reviewers were very specific in their comments on this topic and gave some checks procedures more credits than others (See the subtopics of Measurements in Appendix 1).

While CMT claimed no authority for competence in their committee on interview procedures, the reviewers, as field experts, made about as many observations on the Interview as on the Sampling. They showed concern about the lack of documentation on coding, wording of questions, and checks.

In the topic of Sampling, the concerns are multiple: the size and structure of the sample never clearly specified the overrepresentation of some segments, the defective definition of sampled population, etc. Notice that Wallis, the psychologist/economist/statistician, points out more than one third of the critical observations on Data Presentation. He is particularly concerned about lack of clarity, hard to understand information in tables, contradictory or inconsistent information.[18]

The commentators insisted on the problem of volunteering, the repeated use of one same history to produce various cases, the problem of memory, insufficient or undocumented checking procedures, the non-structured interview and the unclear specifications of the "100% groups" that Kinsey used as

[17] Kinsey is not to be criticized for not using the methods common in public opinion polls; as he points out, a strictly random selection of subjects in a study of sexual behavior would not have been feasible. The report is open to criticisms, for not given us the information needed to judge the representativeness of either the volunteers or the 100% groups. The number of contributing groups is almost never stated." (...). Whatever the number of these individual groups may have been, it is unlikely that the total hundred percent samples could have been representative of the U.S. population, or could have been doctoring." (Terman's review paper)

[18] "My strongest recommendation about future reports is that they should say precisely what was done. My strongest complain about the volume is that when I study it in any detail I frequently cannot tell what information Kinsey's conclusions are really based on."(Wallis Review)

clusters.[19] The other hard critics had doubts about representativeness and "who" were really the Kinsey's subjects or sources of histories and cases.[20]

The inclusion of institutionalized subjects, the long, invasive, extended, and bond creating interview, and the inclusion of referrals and external sources is a source of some comments from the critics. However, the counting of sexual acts as indicator of "sexuality" did not constitute a major concern for them.

In spite of being quite specific in their technical observations, they never considered the ethics of the use of theses histories. When objecting to representativeness of the sample, they also occasionally mentioned the absurdity of extending Kinsey's findings to the U.S male population as a whole. In other words, they showed complete disregard toward the individuals actually represented in Kinsey's research and also toward those, possible the silent majority, not in Kinsey's "sampled population."

It is interesting to notice that out of 123 comments, CMT agrees with the reviewers in 105 cases (84%). However, in more than half of them the committee introduces qualifications, which frequently diminish or at least soften the criticism.[21] Chart 2 summarizes the comments as well as the respective CMT qualification.

[19] "Of the 62 one-hundred groups, 42 were of college level, and 7 were delinquents or inmates of penal and mental institutions. 'Perhaps half" of the stories were obtained through contacts resulting from lectures. 17 penal or correctional institutions provided histories. (…). There are data on 1200 persons convicted of sex offenses. (Wallis review).

[20] Undoubtedly, the crucial question is one of representativeness. One would suspect that those who would, for example attend a lecture on sex would represent a selected audience, and in addition those who would volunteer out of that audience out of that audience would introduce further selectivity.

[21] Table 1 provides just a summary of 123 comments. The interested reader should read the complete classification of comments made by Mr. Romy in CMT report on pages 44 to 150.

Chart 2 - Summary of Comments of CMT Reviewers and their final Qualifications

Item	Comments	CTM Final Qualification	Number of comments in CMT report
Sampling	No enough information on selection techniques and composition of the sample.	Agree. It is practically impossible to define the sampled population.	15
Information	Critics of negative nature.	Agree. The subject of interviewing has not been investigated presently.	14
Report	Questions about memory bias and /or social group influence.	Agree. The two points are very relevant. The critics do not propose any method to study these matters.	5
Stability	Lack of test of significance	Agree. But to develop valid test methods would have been extremely difficult.	6
Measurement	Limitation in check procedures and scope.	Agree. Additionally, smoothing trends do not assure accuracy.	26
Statistical technique	Serious errors and absence of professional statistical advice.	Agree. High degree statistical skills would be needed.	9
Presentation of data	Exasperating lack of clarity. Inconsistencies between tables. Changing totals. Poor information about the content of the interview	Agree. Although recognizes that a complete info on interviews is not an easy requirement to meet.	28
Interpretation	Many interesting and provocative conclusions although not solidly based on data.	Agree. The writing of the book falls below the level of good scientific writing.	20
		Total	**123**

The Comparison with other studies

William O. Jenkins was in charge of working on and reporting the comparison of the KPM report with other studies of reported sexual behavior. The list of books to review was provided by the CMT committee and they were selected from a list suggested by Dr. Kinsey. The writer, as Mr. Jenkins is here referred to, added one book (by E. J. Farris) to that list, although it was not included in the final ranking as presented in Chart 3.

Chart 3 - Comparison of KPM with other sex studies

Research and Publication Information	Title	Rank Given by CMT
Bramley, Dorothy C. Britten, Florence H. 1936. New York: Harper	*Youth and Sex.*	7
Clavis, Catherine B. 1929. New York: Harper	*Factors in the Sex Life of Twenty Two Hundred Women*	3
Dikinson, R. L. and Beam, Lure. A. 1931. Baltimore: Williams and Wilkins	*A Thousand Marriages*	6
Dikinson, R. L. and Beam, Lure. A. 1934. Baltimore: Williams and Wilkins	*A Single Woman*	6
Farris, E. J. 1950. White Plains, N. Y. : The Author Press	*Human Fertility and of the Male*	Not included in Kinsey's list
Hamilton, G. V. 1929. New York: Bani	*A Research in Marriage*	4
Kinsey, Pomeroy, and Martin. 1948. Philadelphia: Saunders	*Sexual Behavior in the Human Male*	1
Landis, C. and Balles M. Marjorie. 1942. New York: Harper	*Personality and Sexuality of the Physically Handicap Woman*	5
Landis, C. et al. 1940. New York: Harper	*Sex in Development*	5
Terman, L. M. et al. 1938. New York: Mc Graw Hill	*Psychological Factors in Marital Happiness*	2

Source: CMT, Appendix B

In Appendix 2 , the works are listed with a summary of the purpose of the study, the author's qualifications, the sample and sampling methods, the interview procedures, and the sampled population (according to the respective author).

The writer ranked the KPM report first, Terman's second and Davis' third. All the rest obtained lower ranks. Both Terman and Davis preceded Kinsey in the campaign for sex research, although possibly for different interests. Terman, who was also one of the most critical commentators of the KPM report, was a pioneer in sex research. More than a decade before Kinsey, he had reported on male sexual behavior, and had collected detailed information on sex activity in marriage.[22] Katerine

[22] Lewis M. Terman (1877-1956). Ph.D. in Psychology from Clark University. His first major undertaking was the revision of the Binet-Simon Intelligence Scale, which he published in 1916. In 1920 he began his longitudinal study of some 1,500 intellectually gifted children, a study that has provided the primary source of today's detailed knowledge of the development of persons with high intelligence. He became curious about marriage as a psychological phenomenon, and conducted an extensive study of several hundred marital couples. His findings, published in 1938, contributed notably to our empirical knowledge of the social psychological aspects of marriage in American society. More than a decade before Kinsey reported on male sexual behavior, Terman had collected detailed information on sex activity in marriage.

Davis, one of the first women to have a PhD in Social Work, had also worked on sex research a generation before Kinsey.[23]

It should be pointed out that Robert Latou Dickinson (1861-1950), ranked sixth by Jenkins; an MD from Long Island College Hospital and a medical doctor was the most significant figure in American sex research before Alfred Kinsey. Almost from the beginning of his practice, Dickinson began collecting sexual histories from his patients. Dickinson marked down certain sexual observations. He was strongly convinced that many difficulties his patients reported -including insomnia, menstrual irregularities, and certain types of pain- had their roots in sexual . The value of his data may be reduced especially to documentation.[24]

The writer concludes that the KPM was superior compared to the others, mainly because of its broader scope of variables and a larger sample of subjects. He recognizes that some of the works are not strictly comparable to KPM, given differences in research purposes. He points out the need for more research regarding the choice between questionnaire and interview as measurement instrument. The writer does not make any reference to the marked differences in research objectives, research environment and professional formation of the authors. He does not analyze any point related to the ethical aspects of the researches.

In Appendix 2 there is a summary of the research methodology and other characteristics of the ten compared books.

[23] Katharine Bement Davis (1860; 1935), social scientist, prison reformer. A member of the first generation of American women to earn doctoral degrees in social science, Davis pioneered the sexual survey a generation before Alfred Kinsey. Her landmark study, *Factors in the Sex Life of Twenty-Two Hundred Women* (1929), provided the earliest nonjudgmental account of same-sex relations among women.

[24] "For the psychologist this study is of very limited value. Collected, from the psychologist's point of view, by an untrained observer, the data is very haphazard and will hardly help either to develop old or stimulate new theoretical work. But what it lacks in depth, it perhaps makes up in numbers, for we have here records of such a mass of cases as, through the very nature of his work, no psychologist could collect. This fact gives the book a utility as a check and a source of corroborative evidence. (Yates, 1937)

IV. The Statisticians Evaluation

"A random selection of three people would have been better than a group of 300 chosen by Mr. Kinsey".
John W. Tukey, Princeton University

The ASA committee concluded basically that the KPM work was the best in its field although it shared some unavoidable stresses in the use of samples principles, measurement validity and inference, with the others.

According to the CMT the research procedures selected by KPM consisted basically of a study of sexual acts by US white male individuals, by a non-probabilistic sample of clusters (mainly institutions), and other individuals contacted by specific means (conference attendants, personal or professional referrals, etc.).

The instrument of measurement was a personal interview. The individuals volunteered their participation. For a summary of these and other specific methodological choices by KPM, as well as some notes and comments by CMT, the material is presented in Chart 4.

Chart 4-KPM Methodological Choices

TOPIC	STATITICAL CONCEPT INVOLVED	CHOICE	EVALUATION	COMMENT
What sort of behavior?	Dependent variable	Choice: orgasm as the central sort of sexual behavior.	This choice is not a matter of general quantitative methodology, and hence falls outside the scope of this committee's task	Improper. The definition of "what to observe" is an essential part of a statistical study.
Whose behavior?	Experimental Unit	White male in US	Definition not clear. If Compared with 1940' census would be concentrated in Indiana and a disproportionate "willing to participate". These comments are not a criticism and the choice is taken as a fact.	Improper. The choice in this topic should indicate the appropriate target population and the available sample population. Both should be a close as possible if the purpose of the study is to be achieved. The difference between both determines the inferential capability of the study.
Observed, recorded, or reported behavior?	Response	Reported behavior	In fact a difference between actual and reported behavior may lead to systematic errors. However CMT finds the choice superior to any other, based on the need for generalizations.	Undefined. This choice should be decided in combination with the measurement instrument (questionnaire)
Interview or questionnaire, and types thereof? Interview oral or questionnaire: The former allows to include more subjects and variations during interviews	Measurement instrument	Interview	The CMT does not pronounce on this topic but recognizes that this choice has further effects. In fact the interview seems more a clinical interview that a psychosomatic test	Improper. In statistical terms, the entire experiment requires possibility of replications. The study was presented as a statistical analysis not as the narrative of clinical experience.

TOPIC	STATITICAL CONCEPT INVOLVED	CHOICE	EVALUATION	COMMENT
Which subject?: 1)Individuals at a time or clusters; 2) demographic segments proportional to population structure or of equal size; 3) catch-as-catch-can basis or some probabilistic scheme	Sample unit	Individual in Clusters; demographic characteristic very similar between subjects; not probabilistic design or selection. "KPM's sample was, in the main a cluster sample, since they built up their sample from groups of people rather than from individuals."	No particular observation is made about this choice though CYM affirms: "KPM's sample was, in the main, a cluster sample, since they built up their sample from groups of people rather than from individuals."	Improper. The choice of unit by KPM (clusters or demographic segments) was not a matter of design but based on subject availability. In any case the absence of any probabilistic scheme makes the distinction irrelevant.
Sample	Sample selection	Group of people	Given The U.S. white male was the target population, our conclusion are that: 1) KPM starting with a non-probabilistic sample was justified…"	
What methodological checks? Choices: type of check and number of checks	Evaluation	The checks included: 1) take-retake, 2) husband -wife, 3)duplicate recording interview, 4) overall comparison of interviews,5) others	3) Was done in an unknown, presumably small case. Not report presented, and they may have been done for training not evaluation purposes.	Confusing. The checks should be a statistical evaluation devise or a training practice, not both at the same time.
How analyzed and presented?	Inference and report	KPM chose to report both row and "US corrections" and make simple comparisons. No detailed computations shown. No use of scale or composite scare was attempted.	Except for those cases, most of the analysis is presented as "straightforward" descriptions. There are some references to standard errors, although they have a limited value since there was not probabilistic scheme. This should not be taking as a criticism, since an accurate indication of significance would have been difficult.	This part of the evaluation expresses rather clearly that it is not a successful inferential work. But once more the tone is misleading and did not definitely explains the scope and validity of the KPM report considered from a statistical point of view

TOPIC	STATITICAL CONCEPT INVOLVED	CHOICE	EVALUATION	COMMENT
How interpreted? Choices: 1)Extent of warning about possible differences between reported and actual behavior, 2) extent of warning about possible differences between the sample population and the entire U.S. white male population, 3) extent of warning about possible sampling fluctuations, 4) Extent of verbal discussion not based on evidence presented, 5) certainty with which conclusions were presented.	Conclusions	1) Emphasis indicating that the differences seem small to KPM. 2) Little discussion. 3) Some early warning, not often. 4) Substantial discussion. 5) Presentation as solid certainty.	"In general the observations seem to have been interprets with more fervor than caution, although occasional qualifications may be found". Competence in verbal descriptions. Uncertainties in inferences about the behavior of white males were "brief, insufficiently repeated and sometimes, entirely lacking." Many of the most interesting statements are not based on the tables or any specific evidence	Misleading. Here the evaluation becomes confused and unprofessional mixing "verbal capacity" with valid and well supported statistical report. The evaluator uses the expression "interesting statements" and at the same time recognizes that they are not based on statistical evidence.

Source: *Statistical Problems of the Kinsey Report*. Statistical American Association. 1954. Pages 7 to 47

Regarding the evaluation that the ASA committee requested, the summary of CMT conclusions is presented in Chart 5.

Chart 5 - CMT Evaluation Conclusions Summarized

Topic	Conclusions
Interviewing	Not an ideal method. But we need to know more about interviewing in general.
Checks	KPM are good. They can afford to supply additional checks.
Sampling	KPM starting with a non-probabilistic sample was justified. In the absence of a sampling benchmark, the present results must be regarded as subject to systematic error of unknown magnitude due to selective sampling (via volunteering)
Analysis	Best described a simple and relatively searching. They did not shirk hard work and their summaries were shrewd descriptive comments rather than inferential statements about clearly defined populations.
Interpretation	KPM showed competence in accurate and understandable verbal descriptions. The uncertainty of inferences to the U.S. white male population was brief, insufficient and sometimes lacking. The discussion on systematic errors is careful with few exceptions.
Comparison with other studies	KPM is outstandingly good. Sampling and volunteering are the same. Checks and geographical and social scope is broader in KPM. Only Davis and Ferris are careful in generalizations and warned their readers.
Major controversial findings	All controversial finding are set forth as well- established conclusions. All are subject to: a) differences between reported and actual behavior; and b) non-probabilistic sampling involving volunteering.

Source: CMT, page 37

In this evaluation CMT places the KPM report as a work with

- practically no inferential power, due to no probabilistic, selective sample of volunteers, and
- a very limited descriptive value due to non-measurable errors in reported behavior and findings not based on supporting data.
- comparatively containing more or less the same limitation of other studies and with a better scope and research quality

Perhaps the evaluators hoped that this description should have been sufficient to disqualify KPM reports as nothing more than a curious and intriguing work to call attention to the necessity of better research work in the field of sexuality, but it was not.

In spite of the CMT clarity and relative objectivity, there was another missing aspect in their evaluation. That is the component of a statistical work, which is not completely detached from technical and procedural aspects. Evidently disregarded by the CMT are:

Concern about the individual: The participation of volunteers gives place to doubts about possible direct or indirect cohersion. In fact, in the case of persons in institutions, participation may be related to pressure from people in charge, need to call attention, and the like. In the case of attendants at events and referrals, it is very probable that the individuals were trying to look for

solutions to their own problems. Referral from professionals may compromise the confidentiality of information received in consultation. In the case of sexual acts, abuse of minors should be reported to authorities, for those events are more than simply "cases" in a data set.[25]

Concern about no possible replications: There was no way of perform replications in KPM interviews, in part, because they were never properly documented. Nevertheless, they were praised by CMT and some commentators. They were said to be coded and that interviewers had memorized hundreds of codes. It seems to be the kind of a conversational contact between fraternal partners. These methods seem not only cohersive but also misleading. This type of instrument is impossible to replicate, and possibly was not really replicated in the KPM experiment. Thus, it breaks all the condition for a measurement instrument of data collection in the statistical sense (besides calling again for ethical considerations).[26]

Concern about researchers becoming part of the experiment: The extreme involvement of the interviewer-individual may be a source of non-measurable variability among individuals and among interviewers. This introduces another obstacle for replication, but even something more serious: the researchers becoming part of the experiment. Possible evidence related to this concern is the fact than many important conclusions are presented without data support, which has to be considered as an evidence of pre-conceived ideas about the reasons for behavior or relation of behavior with other variables.

Concern about representativeness: Although the CMT never stated it clearly, the KPM has no representativeness to make conclusion about the behavior of the US male population. Unfortunately, the size and composition of Kinsey's sample has never been ascertained, not even after the commentator scrutinized the report. They were never able to establish the number of histories, number of cases, size and nature of clusters. [27]

---

[25] In the 1977 book *Ethical Issues in Sex Therapy and Research*,[15] Kinsey coauthor (of the Female Report) Dr. Paul Gebhard makes some very frank statements about how the Kinsey team dealt with some of the ethical issues they confronted. Gebhard's comments go some way toward clarifying the entire Kinsey research philosophy. It was Gebhard's view that "Each researcher must establish his or her own ethical hierarchy and decide as problems present themselves whether the ultimate good resulting from the research or therapy supersedes a particular ethic"[15] (p14). Concerning the nature and sources of information for the Kinsey Reports, Gebhard had this to say: *We have always insisted on maintaining confidentiality, even at the cost of thereby becoming amoral at best and criminal at worst*

[26] "We go with them to dinner, to concerts, to night clubs, to the theater; we become acquainted with them at community dances, in poolrooms, in taverns, and in other places which they frequent. They in turn invite us to meet friends, in their homes, at teas, at dinners, at other social events". (transcript from KPM, appearing in CMT, page 77)

[27] "As to the composition of the sample actually secured, it is hard to learn much about this. Scattered through the book there are various scraps about special groups. Terman, on page 447 of his review, lists some of these: Of the 62 100% groups, 42 were college level, and 7 were delinquents or inmates in penal or mental institutions. Perhaps half of the stories were obtained through contact resulting from lectures. 17 penal or correctional institutions are said to have provided histories. Five 'underworld' communities and five homosexual communities are represented. There are data on 1200 persons convicted for sex offences. In addition, a passage on p.38 strongly suggests to me that several hundred psychoanalysts, psychiatrists,

V. Conclusions

When the CMT report was published, the general public did not have the real dimension of Kinsey's scientific flaws. The CMT evaluation was too obscure or too technical to fully reveal the absurdity of Kinsey's descriptions and generalizations. Nevertheless, the truth was there, in the hands of the statisticians.

It is simply that the smoking gun and the motive of the crime were missing.

- CMT saw the bias in the sample but did not reveal the character of the actual sample.
- When they justified the use of volunteers because it is common in sex research they failed to take a closer look at the reasons for volunteering in the case of KPM.
- When they condescendingly commented on the use of conclusions without supporting data in "many of the most interesting and provocative statements," they failed to search for reasons behind the project.
- When they considered the measure of behavior "out of the statistical scope" they missed the opportunity to look at the lack of conceptualization in Kinsey's work. He looks at human behavior in a similar way to that of a zoologist of his time. He forgets that the human person exists in relation to others and that "others" mean loved ones, family, and community.
- When they considered the lack of representativeness of Kinsey's sample they missed the fact that research work that has the objective of informing public policy has to be representative of all groups.
- When they accepted without analysis that sex research has to be mostly based on volunteers, they failed to understand that in a free society, information is provided by individuals, if some good is perceived to be derived from it. When researchers try to extract information that individuals are reluctant to give, they rely on volunteers; precisely the ones that have reasons to participate. That means that the responsible researcher cannot and must not generalize the results.

Kinsey was in fact a pioneer in sex research. The purpose of his research was to document sexual behavior. His did not care much about statistical rigor, probably because he was convinced of his findings, even before writing his report.[28] He was not interested in the knowhow; he just wanted the

---

physicians, clinical psychologists, social workers, and other professional persons (who) have had an special interest in observing the interviewing techniques were included;…". (From Wallis review)

[28] First, the scientist should not have any preconceived hypothesis in order to present only the facts. "Kinsey actually had a two-pronged hypothesis," Hobbs said. "He vigorously promoted, juggling his figures to do so, a hedonistic, animalistic conception of sexual behavior, while at the same time he consistently denounced all biblical and conventional conceptions of

approval of the highest statistical authorities in the nation, to be accepted in the scientific world; the world of scientific rigor and ethics.

He was a "Scientist" in Oscar Varsavsky's sense: Kinsey, by trying to conquer his place in science, disregarded his obligation to society and even to those who knew less than he knew.

You can fool all the people some of the time. You can fool some of the people all of the time. But you cannot fool all the people all of the time. What we can conclude is that the CMT were scientifically (or politically?) correct in their evaluation and we do not want to say they were unethical. However, in their "correctness" they missed the opportunity to be truthful to individuals and society, at the right time.

This has been a very exhausting and cumbersome task. I have gone through almost all the comments, criticisms, praises and adventures of Dr. Kinsey and his team. I cannot reproduce all the information that is available about this monumental and flawed study. I have included some interesting additional notes in Annex 1 that may complement in part the information in this article.

By the way, one of the Notes is about the NSFG, a survey of the US federal statistical system that provides information on sexual activity and other related matters. It is based on a probabilistic sample; it has a structured questionnaire, a decent response rate and follows all the rules of representativeness, confidentiality, reliability and respect for the individuals that U.S. Title 13 requires from statisticians when they want to ask something from you or me.

---

sexual behavior." Second, Kinsey refused to publish the basic data upon which his conclusions rested. Third, he refused to reveal the questionnaire upon which he based all of his facts. (Susan Brinkmann, CS&T Correspondent. Sordid Science: The Sex Research of Alfred C. Kinsey (The Catholic Standard & Times - Part 2 of 7, August 14, 2005).

References

- American Statistical Association. Committee on Professional Ethics. Ethical Guidelines for Statistical Practice. Executive Summary. 1999
- Anonymous, M.D. Unprotected. Sentinel. Penguin Group: New York. 2008
- Brinkmann, Susan. Sordid Science: The Sex Research of Alfred C. Kinsey (The Catholic Standard & Times - Part 3 of 7, August 14, 2005
- Bromley, Dorothy Dandy Britten, Florence H. New York: Harper. Youth and Sex. 1938.
- Browder, Sue Ellin. "Kinsey's secret: The phony science of the sexual revolution" Crisis, May 2004"
- Cochran, William G., Frederick Mosteller, John H. Tukey and W. O. Jenkins. Statistical Problems of the Kinsey Report on Sexual Behavior in the Human Male. The American Statistical Association. Washington D.C., 1954.
- Davis, Katherine B. Factors in the Sex Life of Twenty Two Hundred Women. New York: Harper.1929.
- Dickinson, R. L. and Beam, Lura - B. A Thousand Marriages. Baltimore: Williams and Wilkins.1931
- Dickinson, R. L. and Beam, Lura - A. The Single Woman. Baltimore: Williams and Wilkins. 1934.
- Farris, E. J. Human Fertility and Problems of the Male. White Plains, NY: The Author Press. 1950.
- Goldstein J, Pastore N. Sexual Behavior of the American Male: A special review of the Kinsey report. The Journal of Psychology. Vol. 26 (1948)
- Hamilton, G. V.A Research in Marriage. New York: Boni. 1929
- Hayman H, Sheatsley Paul B. The Kinsey Report and Survey Methodology. International Journal of Opinion and Attitude Research, Vol. 2 1948
- Hobbs A.H., Lambert R.D. An Evaluation of 'Sexual Behavior of the Human Male'. American Journal of Psychiatry, Vol. 104 (1948).
- Houston, Larry. Behavior and not a person, Chapter 3 Alfred Kinsey 1989
- Jeffrey, Linda. Restoring Legal Protection for Women and Children. The State Factor. American Legislative Exchange Council. 2004
- Kinsey, Alfred C., Warren B. Pomeroy, and Clyde E. Martin. Sexual Behavior in the Human Male. W. B. Saunders Company. Philadelphia and London, 1968.
- Kinsey, Alfred C., Warren B. Pomeroy, and Clyde E. Martin. Sexual Behavior in the Human Male. W. B. Saunders Company. Philadelphia and London, 1968.
- Kinsey, Alfred C., Warren B. Pomeroy, Clyde E. Martin and Paul H. Gebhard. Sexual Behavior in the Human Female. Indiana University Press. Bloomington & Indianapolis, 1998.

- Landis, C. and Bolles, M. Marjorie. Personality and Sexuality of the physically handicap woman. New York: Hoeber. 1942.
- Landis, C., et al. Sex in development. New York: Hoeber.1940.
- Mason, Mervyn L. Mason. Human Sexuality and Critical Evaluation of Recent Texts, Greenwood Press. Wesport: 1983
- Muir Gordon, Court John H. Fraud of the Century? May 1992 Edition Of The Catholic Medical Quarterly
- "Reisman, Judith. Kinsey: Crimes & Consequences. The Institute for Media Education, Crestwood, KY., 1998, 2000."
- Reumann, Miriam G. American Sexual Character. Sex, Gender, and National Identity in the Kinsey Reports. University of California Press. 2006
- Terman, L. M., et all. Psychological Factors in Marital Happiness. New York: McGraw Hill. 1938.
- Terman, L.M. Kinsey's Sexual Behavior of the Human Male"; some comments and criticisms. Psychological Bulletin, Vol. 45, (1948)
- Wallin, P. An appraisal of Some Methodological Aspects of the Kinsey Report. American Sociological Review, Vol. XIV, No.2 (1949).
- Wallis, W.A. Statistics of the Kinsey Report. The Journal of the American Statistical Association, Vol. 44 (1949).
- Yates, S. L. Review of "The Single Woman": By R. L. Dickinson and L. Beam. Williams & Norgate Ltd., and Baillière, Tindall and Cox, London. 1937

Appendix 1 – Additional Notes

Note #1

In spite of its correctness, Dr. Kinsey is not happy with CTM report (Extracted from Larry Huson's "*Alfred Kinsey*", Chapter 3)

> On the whole, the CTM report was very favorable. The writers declared that statistical and methodological aspects of the work were outstanding in comparison with other leading sex studies, and termed the interviewing "of the best". They criticized for not indicating what statements in the book were undocumented or un-documentable and declared that he should have been more cautious in drawing precise conclusion from a limited sample. Some of the findings were questionable for possible bias in the constitution of the sample, but the writers noted that no previous study of any kind had been able to avoid that difficulty and bias could not even be reduced by a probability sampling program.
>
> As for the statistical help we had used, the writers noticed that it had been limited in part because wartime, and concluded that the kind of assistance that might have solved some of the most complex problems would require a knowledge possessed by perhaps no more than twenty statisticians in the world. They went then on to recommend a probability sampling program for future work and made other technical suggestions.
>
> From a letter to George Gallup, where Kinsey wrote:
> Our conference was a consider success from our standpoint, thanks in part to the help which your group gave us. I think the statistical group has not really intended to say some of the things theirs words really meant, and they seem amazingly unaware of the public relation problem that such a report, by its inept phrasing, would present. We understand each other better now , even they would not back down on their insistence that probability sampling was the only perfect thing , and they very well understand that we do not intent to engage in such program , and that we will explain to the world why.

Note #2

Kinseyism? (From "*Kinsey's Secret: The Phony Science of the Sexual Revolution*" by Sue Ellin Browder)

> "In a 1951 Journal of Social Psychology study, psychology students at the University of California, Los Angeles, were divided into three groups: Some students took an intensive nine-week course on Kinsey's findings, while the other two groups received no formal Kinsey instruction.

Afterward, the students took a quiz testing their attitudes about sex. Compared with those who received no Kinsey training, those steeped in Kinseyism were seven times as likely to view premarital sex more favorably than they did before and twice as likely to look more favorably on adultery. After Kinsey, the percentage of students open to a homosexual experience soared from 0 to 15 percent. Students taught Kinseyism were also less likely to let religion influence their sexual behavior and less apt to follow sexual rules taught by their parents."

Note #3

You cannot fool everybody all the time. (From the review of the film *Kinsey* by Chad Wilks. Journal of Religion and film. Vol. 9, No. 1 April 2005)

The movie is honest about the fact that Kinsey's scientific moralism had its own cost. The pain inflicted on Clara by Kinsey's initial infidelity with same-sex research assistant, Clyde Martin (Peter Sarsgaard), is masterfully portrayed. Moreover, Martin nearly loses his own wife when she falls in love with another man, a colleague in the project with whom she has been consensually involved. Finally, the dogmatism of Kinsey's scientism comes across as occasionally heartless, not at all unlike his own father's religious dogmatism.

Note: There exists another film made in 1998 and entitled "*Secret History: Kinsey's Pedophiles*." It aired in England but was never shown in the United States.

Note #4

You cannot fool everybody all the time.

In reality, Kinsey's reports never applied to average people in the general population. In fact, many of the men Kinsey surveyed were actually prison inmates. Wardell B. Pomeroy, Kinsey co-author and an eyewitness to the research, wrote that by 1946 the team had taken sexual histories from about 1,400 imprisoned sex offenders.

Kinsey never revealed how many of these criminals were included in his total sample of "about 5,300" white males. But he did admit including "several hundred" male prostitutes. Additionally, at least 317 of Kinsey's male subjects were not even adults, but sexually abused children. About 75 percent of Kinsey's adult male subjects volunteered to give their sexual histories. As Stanford University psychologist Lewis M. Terman observed, volunteers for sex studies are two to four times more sexually active than non-volunteers. In the December 11, 1949, New York Times, W. Allen Wallis, then chairman of the University of Chicago's committee on statistics, dismissed "the entire method of collecting and presenting the

statistics which underlie Dr. Kinsey's conclusions:' Wallis noted, "There are six major aspects of any statistical research, and Kinsey fails on four."

In short, Kinsey's team researched the most exotic sexual behavior in America -taking hundreds if not thousands of case histories from sexual deviants- and then passed off the behavior as sexually "normal," "natural;" and "average" (and hence socially and morally acceptable).

Note #5

The mystery of tables 31-34 (From Sordid Science: By Susan Brikman. *The Sex Research of Alfred C. Kinsey* (The Catholic Standard & Times - Part 3 of 7) - August 14, 2005)

Some of the most vile sets of statistics came from the infamous Table 34, "Examples of Multiple Orgasm in Pre-Adolescent Males," that appeared in Alfred C. Kinsey's first book. This was the research conducted on children under the age of 13 and presented to the world as proof that erotic arousal was possible in children as young as two months.

"Table 34 was truly grotesque", writes Dr. Judith Reisman in her book, "Kinsey: Crimes and Consequences." "It reported around-the-clock experimental data on infants and young boys. The Kinsey team seemed perfectly at ease when describing the extraordinary data: 'Even the youngest males, as young as two months of age, are capable of such repeated reactions. Typical cases are shown in Table 34. The maximum observed was 26 climaxes in 24 hours (in a 4 year old and a 13 year old) . . .' How was this data collected?"

Three of Kinsey's books were reprinted at the same time, in 1998, to celebrate the 50th anniversary of the publication of Sexual Behavior in the Human Male. Of interest, printed in only one, Sexual Behavior in the Human Female was a new introduction by John Bancraft the current director of the Kinsey Institute for Sex Research. This introduction included a section about the information that was originally presented in Chapter 5 of Sexual Behavior in the Human Male. It was this chapter that Kinsey included information about infant and young male child sexual behavior.

"I decided to check on the sources of this information and found that, without any doubt, all of the information reported in Tables 31-34 came from the carefully documented records of one man. From 1917 until the time that Kinsey interviewed him in the mid-1940s, this man kept notes on a vast array of sexual experiences, involving not only children but adults of both sexes. Kinsey was clearly impressed with by the systematic way he kept his records, and regarded them as of considerable scientific interest. Clearly, his description in the book of the source of this data was misleading, in that he implied that it had come from several men

rather than one, although it is likely that information elsewhere in this chapter, on the descriptions of different types of orgasms, was obtained in part from some of these other nine men. I do not know why Kinsey was unclear on this point; it was obviously not to conceal the origin of the information from criminal sexual involvement with children, because that was already quite clear. Maybe it was conceal the single source which otherwise might have attracted attention to this one man with possible demands for his identification (demands which now have occurred even though he is long dead). It would be typical of Kinsey to be more concerned about protecting the anonymity of his research subjects (and convincing the reader of the scientific value of the information) than protecting himself from the allegations that eventually followed." (Kinsey, Pomeroy, Martin and Gebhard, Sexual Behavior in the Human Female)

"Both Jones and Garthorne-Hardy point out the data was mostly dependent upon the notes taken by a pedophile although Kinsey tried to cover this up by attributing it to varying sources." (Bullough, "*The Kinsey Biographies.*" p.22)

Note #6

Was it the media to blame? (Paragraph from American Sexual Character. Sex, Gender, and National Identity in the Kinsey Reports by Miriam G. Reumann)

"A crucial development shaping postwar culture was the rise of a national mass media and a public receptive to its claims. As the number and variety of media outlets expanded after the war, so too did the parameters of allowable news. Popular magazines openly discussed homosexuality and sexual techniques, and experts complained that Americans were obsessed with sex. The psychiatrist Albert Ellis lamented the "average" citizen's ignorance in his book The American Sexual Tragedy, while the conservative Reader's Digest issued the plaintive query "Must we change our sex standards?" Articles in women's magazines counseled, "We must face the facts about sex," asked "Do Americans commercialize sex?", and encouraged readers to "check your sex I.Q." As the historian Joanne Meyerowitz argues, the media's discussions of Americans' sexual behavior sometimes "expanded the process by which some readers identified new options for themselves in the popular culture."

Note #7

Willingness to Volunteering

The psychology Abraham Maslow (1908-1970), from Brandeis commented on the work of Alfred Kinsey and noted the bias in Kinsey's studies. Maslow had several years earlier written

about the volunteer error ((Journal of Social Psychology, Vol. 16, pp. 259-294, 1942, as cited by Reisman and Eichel, p. 182).

Independently of Maslow, Lewis M. Terman of Stanford University critiqued Kinsey's report in 1948, collaborating with statistician Quinn McNemar. The internal evidence within Kinsey's reported data alone demonstrated to McNemar that there was serious bias. According to Terman, McNemar's calculations "confirm the suspicion that willingness to volunteer is associated with greater than average sexual activity. And since the volunteers account for about three-fourths of the 5,300 males reported in this volume, it follows that Kinsey's figures, in all probability, give an exaggerated notion of the amount of sexual activity in the general population" (L.M. Terman, Psychological Bulletin, 45: 443-459, 1948, as cited by Reisman and Eichel, pp. 20-21). Terman also noted that many of the volunteers came looking for advice on their personal sexual problems, such as learning more about the potential harmful effects of excessive sex. Those volunteering out of a need for advice on such matters are likely to be greatly over represented relative to the general population. Careful random sampling is required if results are to be extrapolated to the general population.

Even apart from the bias introduced by relying mainly on volunteers is the bias introduced by Kinsey's questioning. Rather than devising an objective means of polling people, Kinsey used a "burden of denial" technique which put pressure on his subjects to confess to high levels of sexual activity. Kinsey described this technique himself:

The interviewer should not make it easy for a subject to deny his participation in any form of sexual activity... We always assume that everyone has engaged in every type of activity. Consequently, we begin by asking when they first engaged in such activity. [Sexual Behavior in the Human Male, p. 53, emphasis in original.]

Note #8

QUESTION: How did Kinsey become interested in sexology? (From New River Media Interview with: James H. Jones, Professor of History, University of Arkansas. Author, Kinsey: A Public/Private Life)

James Jones: After Kinsey became world-famous as a sex researcher, the explanation that he gave the public as to how and why he became a sex researcher pointed to something called a marriage course. This course was offered at Indiana University beginning in 1938. It was a time when marriage classes were becoming very popular in colleges and universities across the country. Indiana didn't have one, and he led a group of academics who put a course together.

But it's really, I think, not accurate to say that his interest in sex research begins with this course. Kinsey had really mastered much of the sexology literature and the marriage counseling literature in the late 1920s, early 1930s. I think he read his material largely to learn about himself, to become a more successful or skillful lover with his wife, and also to explore again questions about his own sexuality that he wanted to see what the scientific literature had to say about.

When he starts the marriage course in 1938, he is from the beginning planning to use the course to do sex research. He has put together a questionnaire in anticipation of having students whose histories he can take. He has practiced his technique as an interviewer. He really wants to do things. He wants to disseminate information in that course which he is teaching, but he also wants to be a researcher and he wants to use the students in that course to begin to put together a fuller portrait and understanding of human sexual behavior.

Kinsey has a magpie's love of facts. He is to his fingertips an empiricist and he worships data. In his view, if you can find data, if you can just compile enough information, you can start to have information that casts light on issues that society knows nothing about. His big objection to the literature of the day on sex is that it's morally tainted. And by tainted he means it's proscriptive. It spends a great deal of time making value judgments about what people ought to do. What Kinsey wants instead, is an approach to sexuality that first of all finds out what people actually do, and then instead of prescribing how people should behave, he wants to turn that formula on its head. What he'd really like to do is take behavior and elevate it to normality, a kind of morality of numbers.

When Kinsey was an entomologist, using taxonomic approach to his discipline to make contributions to the theory of evolution, he was a passionate entomologist, he was a passionate taxonomist. He was an enthralled evolutionist, or evolutionary biologist. And the same kind of passion that he brings to the quest for knowledge and for career advancement as a zoologist he brings to his research on human sexual behavior - with one important exception. Increasing the human drive that is part of the man, increasing the passion that is part of his soul for all research he does is a very strong social warrant. Kinsey is at odds with the way society regulates human sexual behavior, and what he wants to see is a much more encompassing ethic of tolerance that will make a room at the table for lots of different kinds of people who don't fit under the cookie cutters of prescribed morality.

Note# 9

It is measured after all...! [U.S. Census Bureau, (Statistical Abstract of the United States: 2008). National Survey of Family Growth (NSFG)]

Universe, Frequency, and Types of Data: Periodic survey of men and women 15–44 years of age in the household population of the United States. Interviews were conducted in 2002 in person by trained female interviewers. Interview topics covered include births and pregnancies, marriage, divorce, and cohabitation, sexual activity, contraceptive use, and medical care. For men, data on father involvement with children were collected. The most sensitive data—on sexual behavior related to HIV and Sexually Transmitted Disease risk—were collected in a self-administered form in which the data are entered into a computer.

Type of Data Collection Operation: In the 2002 (Cycle 6) NSFG, the sample was a multistage area probability sample of men and women 15–44 years of age in the household population of the United States. Only one person 15–44 was selected from households with one or more persons 15–44. Data were collected and entered into laptop (notebook) computers. In the self-administered portion, the respondent entered his or her own answers into the computer. Sample included 12,571 interviews. The response rate was 79 percent. Hispanic and Black persons, as well as those 15–19 years of age, were sampled at higher rates than White adults. All percentages and other statistics shown for the NSFG are weighted to make national estimates. The weights adjust for the different rates of sampling for each group, and for nonresponse.

Data Collection and Imputation Procedures: When interviews are received, they are reviewed for consistency and quality, and analysis variables (recodes) are created. Missing data on these recodes were imputed using multiple regression techniques and checked again for consistency. Variables indicating whether a value has been imputed ("imputation flags") are included on the data file.

Estimates of Sampling Error: Sampling error codes are included on the data file so that users can estimate sampling errors for their own analyses. Sampling error estimates for nine illustrative analyses are shown on the NSFG Web site at <http://www.cdc.gov/nchs /nsfg.htm>. Sampling error estimates are also shown in most NCHS reports.

Other (non-sampling) Errors: In any survey, errors can occur because the respondent (the person being interviewed) does not recall the specific fact or event being asked about. The NSFG questionnaire in 2002 was programmed to check the consistency of many variables during the interview, so that the interviewer and respondent had a chance to correct any inconsistent information. Further checking occurred after the interview and during recoding and imputation. Typically, less than 1 percent of cases need imputation because of missing data.

Note #10

ASA social ethics for statisticians (From the Executive Summary of Ethical Guidelines for Statistical Practice. American Statistical Association Prepared by the Committee on Professional Ethics. Approved by the Board of Directors, August 7, 1999 - Statistics and Society)

> The professional performance of statistical analyses is essential to many aspects of society. The use of statistics in medical diagnoses and biomedical research may affect whether individuals live or die, whether their health is protected or jeopardized, and whether medical science advances or gets sidetracked. Life, death, and health, as well as efficiency, may be at stake in statistical analyses of occupational, environmental, or transportation safety. Early detection and control of new or recurrent infectious diseases depend on sound epidemiological statistics. Mental and social health may be at stake in psychological and sociological applications of statistical analysis.
>
> Effective functioning of the economy depends on the availability of reliable, timely, and properly interpreted economic data. The profitability of individual firms depends in part on their quality control and their market research, both of which should rely on statistical methods. Agricultural productivity benefits greatly from statistically sound applications to research and output reporting. Governmental policy decisions regarding public health, criminal justice, social equity, education, the environment, the siting of critical facilities and other matters depend in part on sound statistics.
>
> Scientific and engineering research in all disciplines requires the careful design and analysis of experiments and observations. To the extent that uncertainty and measurement error are involved -as they are in most research- research design, data quality management, analysis, and interpretation are all crucially dependent on statistical concepts and methods. Even in theory, much of science and engineering involves natural variability. Variability, whether great or small, must be carefully examined both for random error and for possible researcher bias or wishful thinking.
>
> Statistical tools and methods, like many other technologies, can be employed either for social good or for evil. The professionalism encouraged by these guidelines is predicated on their use in socially responsible pursuits by morally responsible societies, governments, and employers. Where the end purpose of a statistical application is itself morally reprehensible, statistical professionalism ceases to have ethical worth.

Note #11

ADMISSION (Human Events, August 18, 2003)

Kinsey Institute researcher William Simon confirmed "Kinsey interviewed 18,000 people and used only a quarter of the cases in his two reports," meaning data from only about 4,500 total males and females were actually used for the Kinsey studies. W. Allen Wallis, University of Chicago statistician and former president of the American Statistical Association, said Kinsey's data reveal he did not report on those he said he interviewed.

Simon's admission that Kinsey secretly dumped three-quarters of his data coincides with the admission by Kinsey co-author and bogus statistician Clyde Martin. In a private letter (dated Dec. 13, 1990) to Kinsey Institute Director June Reinisch, Martin confessed that criminals and homosexuals and other atypical men were mislabeled in Kinsey's alleged "report" as representing the sexual behavior of your average "human male." Thus, aberrant men were the ALIMPC statistics that changed American sex laws.

Appendix 2 - Summary of Comments by KPM Reviewers

ID	Item	Commentator	Detail	Sub-Item	CMT Comment(*)
1	Data Presentation	Goldstein and Pastore	Not one table showed a total of 5300 cases alleged	Male interviews	AQ
2	Data Presentation	Wallis	The number 5300 white males is cot confirmed or any table	Number of Interviews	AQ
3	Data Presentation	Wallis	Discrepancies in the numbers in the map. Cases and Histories have different meanings	Number of cases	AQ
4	Data Presentation	Wallis	It is not quite clear that the total number of histories is 12,214 as alleged	Number of Observations	AQ
5	Data Presentation	Goldstein and Pastore	Misinterpretation of results	Accumulative Incidence Curves	D
6	Data Presentation	Goldstein and Pastore	One individual used as many cases	Accumulative Incidence Curves	AQ
7	Data Presentation	Goldstein and Pastore	Presentation gives inflated idea of incidence	Accumulative Incidence Curves	D
8	Data Presentation	Goldstein and Pastore	Unspecified amount of overlapping in age categories	Accumulative Incidence Curves	A
9	Data Presentation	Goldstein and Pastore	Lack of consistency in terminology	Inconsistencies	A
10	Data Presentation	Goldstein and Pastore	Unclear description of interview procedure	Interviewing procedures	AQ
11	Data Presentation	Goldstein and Pastore	No Information on characteristics of non-respondents by items	Interviewing procedures	A

ID	Item	Commentator	Detail	Sub-Item	CMT Comment(*)
12	Data Presentation	Goldstein and Pastore	Expansion not explained	US Corrections	A
13	Data Presentation	Hobbes and Lambert	Use of a single story to provide "cases"	Accumulative Incidence Curves	AQ
14	Data Presentation	Hobbes and Lambert	Dubious means of describing behavior	Accumulative Incidence Curves	AQ
15	Data Presentation	Hobbes and Lambert	Not enough degree of representativeness	Accumulative Incidence Curves	D
16	Data Presentation	Terman	Omission of wording of questions	Interviewing procedures	AQ
17	Data Presentation	Wallin	Omission of wording of questions	Interviewing procedures	AQ
18	Data Presentation	Wallis	Confusing tables	Content of tables	A
19	Data Presentation	Wallis	Many do the 162 tables do not admit a clear interpretation	Content of tables	A
20	Data Presentation	Wallis	Tables hard to understand and with discrepancies	Content of tables	AQ
21	Data Presentation	Wallis	Lack of consistency in terminology	Inconsistencies	A
22	Data Presentation	Wallis	No information of Non response by item	Interviewing procedures	AQ
23	Data Presentation	Wallis	Omission of wording of questions	Interviewing procedures	AQ
24	Data Presentation	Wallis	Tables hard to understand and with discrepancies	Interviewing procedures	AQ
25	Data Presentation	Wallis	Contradictory information regarding number of cases	Interviewing procedures	AQ
26	Data Presentation	Wallis	Contradictory information regarding number of cases	Interviewing procedures	AQ
27	Data Presentation	Wallis	Impossible to verify	US Corrections	A

ID	Item	Commentator	Detail	Sub-Item	CMT Comment(*)
28	Data Presentation	Wallis	Unable to understand the expansion	US Corrections	A
29	Interpretation	Goldstein and Pastore	Differential of stability in different ages groups	Stability of sexual patterns	D
30	Interpretation	Goldstein and Pastore	Defective Data	Vertical Mobility	D
31	Interpretation	Goldstein and Pastore	Invalid generalizations	Vertical Mobility	D
32	Interpretation	Goldstein and Pastore	Statistical validity hard to evaluate	Vertical Mobility	A
33	Interpretation	Hobbes and Lambert	Careless in reporting	Homosexuality	D
34	Interpretation	Hobbes and Lambert	Conclusions not in agreement with data	Homosexuality	D
35	Interpretation	Hobbes and Lambert	Not representative sample	Homosexuality	AQ
36	Interpretation	Hobbes and Lambert	Not Valid comparison between older and younger generations	Stability of sexual patterns	DQ
37	Interpretation	Terman	Very low representation of some of sex/age segments	Generalizations based on too small tables	A
38	Interpretation	Terman	Not founded inferences	Generalizations based on too small tables	A
39	Interpretation	Terman	Judgments of evaluation or interpretation for which no data or only inadequate data, are given.	Non supporting data	A
40	Interpretation	Terman	Data not really belonging to two different generations	Stability of sexual patterns	A

ID	Item	Commentator	Detail	Sub-Item	CMT Comment(*)
41	Interpretation	Terman	Conclusions based on insufficient evidence regarding marital status	Vertical Mobility	A
42	Interpretation	Terman	Conclusions based on insufficient evidence regarding religion group	Vertical Mobility	A
43	Interpretation	Terman	Defective Data	Vertical Mobility	A
44	Interpretation	Terman	Defective Data	Vertical Mobility	D
45	Interpretation	Wallin	Impossibility of relate conclusion to specific data	Non supporting data	AQ
46	Interpretation	Wallis	Many assertions are of considerable interest but not investigated to be supported	Non supporting data	AQ
47	Interpretation	Wallis	Not clear distinction between methodological, anthropological and reliability type of information	Non supporting data	A
48	Interpretation	Walllis	Defective Data	Vertical Mobility	A
49	Interview Problem	Goldstein and Pastore	Way of asking may have induced serious biases	Interviewing technique	AQ
50	Interview Problem	Hobbs and Lambert	The used technique is very effective	Face to face interview	A
51	Interview Problem	Hyman and Seatsley	Contradictory characteristics in the description of relation interviewer-subject	Interviewer respondent relation	A
52	Interview Problem	Hyman and Seatsley	Investigator had to memorize up to 532. Details if coding never reviled	Standardization of questioning	A
53	Interview Problem	Terman	Concepts hard to be define with precision and hard to code	Coding	AQ
54	Interview Problem	Terman	The investigator knowing the identity. It may hurt accuracy	Face to face interview	A
55	Interview Problem	Terman	Burden of denial on the subject. It may introduced bias	Interviewing technique	A
56	Interview Problem	Terman	The report shows little concern about fabrication	Interviewing technique	A

ID	Item	Commentator	Detail	Sub-Item	CMT Comment(*)
57	Interview Problem	Terman	No other investigator can repeat the experiment for lack of standardization and unrevealed coding	Standardization of questioning	A
58	Interview Problem	Wallin	No supported position against the structured questionnaire	Face to face interview	A
59	Interview Problem	Wallin	Leading questions biasing replies	Interviewing technique	AQ
60	Interview Problem	Wallin	Only the judgment of the interviewer could estimate whether they achieved the purpose	Interviewing technique	A
61	Interview Problem	Wallis	No conclusive statement in favor of one technique	Face to face interview	A
62	Interview Problem	Wallis	Report unclear about the detailed procedure.	Standardization of questioning	AQ
63	Measurement Problem	Goldstein and Pastore	Reliability, not validity	Comparison of interviews	D
64	Measurement Problem	Goldstein and Pastore	Validity subject to specific assumptions	Delay Comparisons	A
65	Measurement Problem	Goldstein and Pastore	Confusing units of measurements	Retake, spouse comparisons and the like	AQ
66	Measurement Problem	Goldstein and Pastore	No information of how many were asked for retakes	Retake, spouse comparisons and the like	D
67	Measurement Problem	Goldstein and Pastore	Doubts about smooth trends as evidence of validity	Smooth trends	
68	Measurement Problem	Hyman and Sheatsley	Consistent interviewing	Retake, spouse comparisons and the like	D
69	Measurement Problem	Terman	Reliability, in some limited and specific cases	Comparison of interviews	D
70	Measurement Problem	Terman	Good correlation but not in variable under study	Retake, spouse comparisons and the like	A

ID	Item	Commentator	Detail	Sub-Item	CMT Comment(*)
71	Measurement Problem	Terman	Retakes do not test validity	Retake, spouse comparisons and the like	AQ
72	Measurement Problem	Wallin	The reports does not provide enough information on age at the time of recall	Comparison with observed behavior	A
73	Measurement Problem	Wallin	Not legitimate assumptions	Delay Comparisons	A
74	Measurement Problem	Wallin	Incompleteness and discrepancies in tables	Retake, spouse comparisons and the like	AQ
75	Measurement Problem	Wallin	Misleading the lector in conclusions	Retake, spouse comparisons and the like	AQ
76	Measurement Problem	Wallin	No indication about the circumstances under which the subjects were procured for re-take histories	Retake, spouse comparisons and the like	AQ
77	Measurement Problem	Wallin	Tests were made only for some age groups	Retake, spouse comparisons and the like	A
78	Measurement Problem	Wallin	Claim of reasonably small error unjustified	Smooth trends	A
79	Measurement Problem	Wallin	insufficient Indication of trend	Smooth trends	D
80	Measurement Problem	Wallin	Misleading smooth accumulative incidence curves	Smooth trends	A
81	Measurement Problem	Wallin	Not valid inference on accuracy of data	Smooth trends	D
82	Measurement Problem	Wallin	Sample heavily weighted with college level mates	Smooth trends	A
83	Measurement Problem	Wallin	Steady trends with educational level not observed	Smooth trends	AQ
84	Measurement Problem	Wallis	In limited cases, acceptable degree	Comparison of interviews	D
85	Measurement Problem	Wallis	Data not available for a given age group by age at the time of interview	Delay Comparisons	A

ID	Item	Commentator	Detail	Sub-Item	CMT Comment(*)
86	Measurement Problem	Wallis	Questionable accuracy testing	Retake, spouse comparisons and the like	D
87	Measurement Problem	Wallis	Too erratic to provide confidence to smoothing	Smooth trends	AQ
88	Measurement Problem	Wallis	Use of a single story to provide "cases"	Smooth trends	A
89	Report Problem	Terman	Long distance memory report as a source of error	Influence of time dimension	AQ
90	Report Problem	Wallin	Attitude may be important to invalidate data	Influence of attitude dimension	AQ
91	Report Problem	Wallin	Important limitations. Average hard to estimate for some subjects	Influence of time dimension	AQ
92	Report Problem	Wallin	Report fails to give age at the time of recall.	Influence of time dimension	AQ
93	Report Problem	Wallin	Report fails to give good reasons to justify combining data for old mad young men for the same age	Influence of time dimension	AQ
94	Sampling Problem	Goldstein and Pastore	It would have been better to start with the state of Indiana and do some stratification	Who are in the sample	AQ
95	Sampling Problem	Hyman and Seatsley	The %100 groups may not be representative	Volunteers	AQ
96	Sampling Problem	Hyman and Seatsley	Nowhere is the description of the characteristics of the 5300 respondents actually interviewed, the number of cases in major cells and the number of refusals	Who are in the sample	AQ
97	Sampling Problem	Terman	Inadequacies of the sample , the attempted "corrections " to sow hypothetical incidences and frequency for US population is indefensible	Other sources of systematic error	A
98	Sampling Problem	Terman	Volunteers account for about three fourth of the 5300 males. This may lead to exaggerate notion of sexual activity	Volunteers	AQ
99	Sampling Problem	Terman	Lacking information to judge representativeness	Who are in the sample	AQ

ID	Item	Commentator	Detail	Sub-Item	CMT Comment(*)
100	Sampling Problem	Terman	5300 white males in the sample. Not clear how many sex offenders are included	Who are in the sample	AQ
101	Sampling Problem	Wallin	A portion of the sample was secured through "contacts". In some cases by being paid	Other sources of systematic error	A
102	Sampling Problem	Wallin	No information of the number of paid subjects	Other sources of systematic error	A
103	Sampling Problem	Wallin	Limited to a region and expanded to the nation	Other sources of systematic error	A
104	Sampling Problem	Wallin	Volunteers may have participated because they were looking for information on their personal	Volunteers	AQ
105	Sampling Problem	Wallin	Disproportionate representation of deviants, prostitute, and feebleminded persons	Who are in the sample	AQ
106	Sampling Problem	Wallis	Necessity of searching for the effective "sampled population " (that is evidently not US male)	Who are in the sample	AQ
107	Sampling Problem	Wallis	Fragmented and incomplete information prevent to have exact sample size	Who are in the sample	AQ
108	Sampling Problem	Wallis	Judgment of representativeness difficult for lack of information on sources, size and composition of subgroups	Who are in the sample	AQ
109	Stability Problem	Wallin	Not proper interpretation and misapprehension about ability of different measures rendered the effort worthless	S	AQ
110	Stability Problem	Wallis	Many of cases represent the same individual	Reporting values based on few cases	AQ
111	Stability	Hyman and Sheatsley	Faces sampling as in quota sampling and tries to solve it	Sample of clusters	D

ID	Item	Commentator	Detail	Sub-Item	CMT Comment(*)
112	Stability	Terman	Method to determine subgroup sample size revels misinterpretation of nature of variability	AQ	AQ
113	Stability	Wallis	Significant testing needed	Need for significance tests	AQ
114	Stability	Wallis	Some is cluster sampling of groups: sororities ,hitch-hikers-hikers, or mental institutions	Sample of clusters	A
115	Statistical Techniques	Goldstein and Pastore	Age and date of the interview	Comparisons (regarding date)	A
116	Statistical Techniques	Terman	Two individuals in different groups do not all belong to separate generations	Comparisons	AQ
117	Statistical Techniques	Wallis	Comparison of a part with the whole	Comparisons	AQ
118	Statistical Techniques	Wallis	Mishandle of range	Definitions of mean, median, etc.	A
119	Statistical Techniques	Wallis	Not use of sequential method	Definitions of mean, median, etc.	D
120	Statistical Techniques	Wallis	Wrong formulas	Definitions of mean, median, etc.	AQ
121	Statistical Techniques	Wallis	Lack of statistical consult	Overall quality of statistical techniques	A
122	Statistical Techniques	Wallis	Multiple use of same histories	Overall quality of statistical techniques	AQ
123	Statistical Techniques	Wallis	Multiple use of same histories	Overall quality of statistical techniques	AQ

Source: CMT, Appendix A.
Note: (*) A: Agrees; AQ: Agrees with qualifications; D: Disagrees

Appendix 3 - Comparison of KPM with other Sex Studies

Part 1 (Bromley through Farris)

Authors	**Bromley, Dorothy D. and Britten, Florence H.**	**Davis, Katherine B. 1929**	**Dickinson, R.L. and Beam, Lura - A. 1931**	**Dickinson, R.L. and Beam, Lura - B. 1934**	**Farris, E.J. 1950 (Not included in Kinsey's list)**
Title	Youth and Sex. 1938.	Factors in the Sex Life of Twenty Two Hundred Women.	A Thousand Marriages.	The Single Woman	Human fertility and Problems of the Male
Stated purpose	"…we have tried to secure the facts about the sexual habits of the younger generation…"	"Because of the lack of data as to normal experience of sex, on which to base educational programs…"	"The general health and circumstances and fertility of wife and husband are studied."	It deals with women. It stresses anatomical and physiological characteristics and changes due to sexual practices.	The understanding of and application of the scientific principles of reproduction.
Sample and Sampling Methods	1364. (5000 questionnaires sent). "We chose college students for the purpose of this study because they are easy to reach in large numbers…"	1000 married women, 1200 single college graduate, 50 married interviews. The two large samples came from a set of 10000 sent by mail to all the nation	1050 cases. Contact: Visit to doctor because a problem in childbirth, pelvis disorder or marital relationship. Individual in professional , well off type of social group	"A total of 1078 records fairly represent the scope and character of the doctor's private practice as a gynecology and obstetrician". The work with two major "control groups", although they role in the analysis is not very clear. The individual belongs to an educated minority. Notice that they were under treatment.	Samples of different sizes were used. They have N's of 100, 478, 643, and 1000. He points out the a-typically of his sample: "..It is comprised largely of 1) husbands whose wife has not conceived and 2) volunteer donors."

Authors	**Bromley, Dorothy D. and Britten, Florence H.**	**Davis, Katherine B. 1929**	**Dickinson, R.L. and Beam, Lura - A. 1931**	**Dickinson, R.L. and Beam, Lura - B. 1934**	**Farris, E.J. 1950 (Not included in Kinsey's list)**
Qualifications of Authors	The authors are not professional sociologists, psychologists or physiologists. They describe themselves as "reporting journalists".	Social worker with a doctoral degrees in social science	The senior author is a specialist on gynecology and obstetrics; the junior author has experience in the field of education	The senior author is a specialist on gynecology and obstetrics; the junior author had experience in the field of education	The author is a PhD in zoology and anatomy
Sample population according to authors	General youth of US.	The author seems to intend her inferences to be restricted to the samples at hand and similar ones: single women with college background and married women with less education that requested the questionnaire.	The authors seem to intend the inferences to be applied to the population of American married couples.	The authors seems to intend the inferences to be applied to the population of post pubertal women	Apparently the author intended his conclusions restricted to his sample: volunteers and males whose wives failed to conceive
Interview	The questionnaire "grew out of talks"	The complete questionnaire is not presented. It is said to be the result of consultation with human behavior specialists	The patient was required to bring to the first appointment 4 page questionnaires on family and general history, particular illness and special symptoms. The interview was not standardized-each patient was treated as an individual.	Not described in detail. Presumably they followed the procedures described for Dickinson and Beam -A.	A private interview is conducted but details are not provided. A personal history card is given. The basics of reproduction are explained to the couple and confidentiality of the responses is assured. ,

Authors	**Bromley, Dorothy D. and Britten, Florence H.**	**Davis, Katherine B. 1929**	**Dickinson, R.L. and Beam, Lura - A. 1931**	**Dickinson, R.L. and Beam, Lura - B. 1934**	**Farris, E.J. 1950 (Not included in Kinsey's list)**
Statistical Method	Means and medians reported. No measures of variability	Means and medians reported. Variability is measured by ranges	Frequency distributions with cumulative number of cases and percent are employed in a few stances. Breakdowns are made in term of age, education, and religion	Frequency distributions and percentages are the only descriptive statistics used, besides the computation of the median in some cases. They wrote an important note: Precise figures and even proportions are not important data here because the material has its own meaning but may not be quantitatively like the world outside."	Frequency distributions and percentages are the only descriptive statistics employed. No use is made of tests of significance and association
Checks	Not complete. They feel that the students "were honest". Catch questions were given but no report is offered on the check procedure and results.	No numerical evidence of internal consistency presented. Little was done to check validity	Physical examination and patient's report were compared and corrected according to the doctor's criteria. Patients were seen frequencies and corrections were done if necessary. Data were collected for 100 husbands but not report is presenting about cross-checking.	The use control groups may be considered a kind of checks. Occasional references are made to agreement between physical examinations and patient' statements. Eleven husbands were interviewed but not analysis is reported.	No information is provided concerning internal checks on reported behavior

Part 2 (Hamilton through Terman)

Authors	**Hamilton, G. V.** **1929**	**Kinsey, Pomeroy, and Martin** **1848**	**Landis, C.** **And** **Bolles, M. Marjorie** **1942**	**Landis, C.** **and** **et all.** **1940**	**Terman, L. M.** **1938**
Title	A Research in Marriage	Sexual behavior in the human male	Personality and Sexuality of the physically handicap woman	Sex in development.	Psychological Factors in Marital Happiness.
Stated purpose	To investigate whether difficulties in marriage were a function of the marriage institution or of behavior learned in childhood.	It is a fact finding survey in which an attempt is being made to discover what people do sexually, and what factors account for differences in sexual behavior among individuals and among various segments of the population.	To answer the question: To what extent do environmental or social conditions affecting the developing of the individual influence adult sexuality.	To answer the 3 questions: 1) What is the normal pattern of psychosexual development? 2) How deviations from this pattern affect the adult personality? 3) What were the characteristics of psychosexual development of different type of adult personality?	To investigate psychological factors in marital happiness.

Authors	**Hamilton, G. V.** **1929**	**Kinsey, Pomeroy, and Martin** **1848**	**Landis, C.** **And** **Bolles, M. Marjorie** **1942**	**Landis, C.** **and** **et all.** **1940**	**Terman, L. M.** **1938**
Sample and Sampling Methods	Mainly composed of volunteers. 100 male and 100 female. 55 couples were represented. However, he draws conclusions to be applied to the US general population	The authors mention 5300 interviews with white males, mainly from the NE part of the nation. Many have participated as volunteers. Payments have been confined to the very poor. The sample is heavily loaded with individuals from special groups such as college students, prisoners. Members of mental institutions, and male prostitutes.	N= 100 cases including four general categories of physical handicap. The subjects were obtained through the cooperation of 12 different agencies. The criterion for selection was age intellectual capacity, natural of handicap, and age at onset of physical handicap. They were aware that the sample may not be representative of the total population of handicap, but believed that the facts and relations found are probably present in any other sample similarly selected. A sample of not handicapped was also selected.	The authors mention 5300 interviews with white males, mainly from the NE part of the nation. Many have participated as volunteers. Payments have been confined to the very poor. The sample is heavily loaded with individuals from special groups such as college students, prisoners. Members of mental institutions, and male prostitutes.	In the pilot study there were 341 married couples and 100 divorced couples, a total of 900 individuals. For the main study there were 792 individual couples. Some were secured by their attending to related conferences, and some were clients of two major family institutions. Quote: "Our population as a whole includes a fairly wide sample of the married population of south California, but it probably contains a rather large proportion of subjects interested in uplift activities or in matters of self-improvement". This bias, doubtlessly contributed to the cooperativeness of the subjects.
Qualifications of Authors	The author is a Psychiatrist	The senior author is a zoologist and the first junior author is a psychologist	Both authors are psychologists	The senior author is a psychologist. Four other psychologist and medical men worked in the study. One was Dickinson.	The senior author is a psychologist.

Authors	**Hamilton, G. V.** **1929**	**Kinsey, Pomeroy, and Martin** **1848**	**Landis, C.** **And** **Bolles, M. Marjorie** **1942**	**Landis, C.** **and** **et all.** **1940**	**Terman, L. M.** **1938**
Sample population according to authors	It seems that the author expected his inference to be applied to U.S. adults	It seems that the author expected his inference to be applied to white males in the U.S.	Intended to infer to the population of adults, physically handicap U. S. female.	Their populations consists of :1)institutionalized U.S. female adults and 2) non-institutionalized U.S. female adults	Term an states that his inferences apply to middle and upper middle class married couples of urban and semi-urban California as of the time the study was done
Interview	There were 357 questions for women and 334 from men. The questionnaire is fully documented. Each subject was examined individually. When couples participated, they were asked not to communicate	"In each history, up to 521 items have been explored. In average interviews took about 2 hours. Answers were coded. Questions and codes are not provided by the study.	Three major indices were used: a controlled interview, the Rorschach, and the medical history. The material in a number of topics was pull together into scales, e.g., early sex information. In most cases the medical history was obtained from the physician of the handicap person. In a few cases that was not possible and the subject has to be interviewed.	Cards were presented one at a time to the subjects and the response was written down verbatim by the interviewer on 58-66 items. With some exceptions, they tried to apply the same questions to both groups	A questionnaire was used. It was the product of a 2 year preparatory work. It contained 400 items, among which only 20 were direct questions about sexual behavior, although some other questions were asked in relation to opinion and attitude. The questionnaires were anonymously distributed in groups of 40 couples and husband and wife filled them up while sitting in opposite sides of the same room. No notes are given about interviewing training.

Authors	**Hamilton, G. V.** **1929**	**Kinsey, Pomeroy, and Martin** **1848**	**Landis, C.** **And** **Bolles, M. Marjorie** **1942**	**Landis, C.** **and** **et all.** **1940**	**Terman, L. M.** **1938**
Statistical Method	Only frequency counts and percentages are computed. Various breakdown are presented (like marital satisfaction index) separately for men and women. This is not presented as statistical standpoint but only as facts of observation, and exploring explanatory possibilities.	The statistical manipulation of the data is said to be kept at the minimum. There are individual and group frequencies, frequency curves, means, medians, and 'cumulative incidence curves", a technique developed by Kinsey. Expansion coefficients for Us population were used. Thus this author seems to have intended inferences to the population of white male in the US	Use of numbers and percent's of the 100 cases and the non-physically handicap were computed. Chi square and "yates" correlation for continuity were computed.	Numbers and percent's were computed .The test of significance consisted of comparing normal with abnormal by applying Chi-square to the frequencies for single women, married women and total	The principal statistical devises are correlation methods and item analysis. A "happiness" score was used and its results are presented in term of the background variables. Critical ratios of percent differences between high and low marital happiness groups and tetrachorie correlations for husband and wife agreement for both levels were presented.
Checks	No checks procedure is presented. Only the fact that they were given opportunity to review the answer and eventually make changes.	Retake of the whole history, mainly for college students. Comparisons of spouses were performed. Other check devise was the comparison between interviewers. Comparisons were made of data obtained in different time periods. Not all of these checks procedures are documented in detail.	Regarding reliability, they expressed that no subject would intentionally distort heir report. Rorschach test was applied although the results are not reported.	Little quantitative information is provided concerning reliability and validity of the several forms. Reference, however, is made to a number of checks such as critical review of the forms, rechecking of questions when reply was vague. No numerical report on these checks is provided.	The principal checks used were comparisons between husband and wife. The results were encouraging. Additionally, it is noted that almost all the subjects answered all the questions.

Source: *Statistical Problems of the Kinsey Report, Pages 153-217*