**A CRITICAL REVIEW OF RESEARCH ON THE HUMAN/COMPANION ANIMAL RELATIONSHIP: 1988 TO 1993**

By: Beth Ellen Barba


***Note: Figures may be missing from this format of the document***

**Abstract:**

Fifty-two research reports from the human/companion animal relationship literature published from 1988 to 1993 were reviewed using a shortened form of the Selby Research, Assessment Form II (RAF). Descriptive data were analyzed for characteristics such as attributes of authors, grant funding, purposes, quality of literature reviews and conceptual frameworks, settings and sampling, research designs, and implications for future research. Qualitative data describing results of studies were summarized. Over half of the studies were published in *Anthrozoös* and were not grant funded. These were primarily nonexperimental studies using nonprobability, nongeneralizable samples. Social support was the most frequently cited framework. Understudied populations were rural groups, the middle aged, adolescents, and institutionalized elders. Few errors were committed in describing instruments or reporting analyses of data.

**Article:**

**INTRODUCTION**

Over the last ten years, the human/companion animal relationship has become recognized as an important area for research. As might be expected, the volume of literature in this area has been growing rapidly. Hence, it is important both to identify and describe existing research related to the specialty and to assess the quality of this research as critique and evaluation of the scientific merit of the research should precede the application of research to practice.

For many years the human/companion animal research literature has been criticized as being methodologically flawed and in particular lacking a clear theoretical basis and composed primarily of unscientific anecdotal reports with few high-quality experimental studies.

This article describes and evaluates 52 research reports on human/companions from 1988 to 1993. The review examines the general characteristics of those published research articles: educational preparation of authors, journals, grant finding, etc.; the purposes, quality of literature reviews, and conceptual frameworks of the research; types of settings and sampling techniques used; types of research designs employed; common problems in the methodology, sampling, and analysis; and stated implications for future research.

**METHODS**

*Sampling Criteria for the Review*

The review focused on the human/companion animal relationship as it is currently portrayed. Therefore, only research published over the past 6 years, January 1988 through September 1993, was examined. Reports of research on the relationships between humans of all ages and companion animals were identified by key words and abstracts. Because the purpose was to evaluate the scientific methods used and gaps in knowledge, only reports that presented findings and analyses of human/companion animal-related data were included. Unpublished research reports were not included. Also, manuscripts not published in English, theoretical discussions, reports of psychometric properties of instruments, and studies of primarily animal behavior were excluded.
**Sampling Procedure**

Four computerized bibliographic databases for the time period of the review were searched: CINAHL, ERIC, Medline, and Psych Abstracts. CINAHL includes literature in nursing and other health fields. ERIC covers educational materials not covered in other databases. Medline focuses on medical and veterinary literature, and Psych Abstracts deals with psychologic and sociologic literature. Searches of the databases produced 69 articles that met the criteria listed above.

A careful manual search of journals known to publish human/companion animal research, various reference lists, and reference citations was then conducted to ensure that the project was as complete as possible. This search yielded 17 additional reports. The final sample consisted of 52 of the 86 articles identified that met the criteria; these 52 were those available to the author. Thus the sample was one of convenience.

**Instruments**

The shortened form of the Selby Research Assessment Form II (RAF) (Selby 1988) guided the evaluation of the articles. The shortened form of the RAF contains 67 items for assessing adherence to methodological standards in research, from the title and purpose of a study through the methods, results, and interpretation of the results.

The RAF has been judged to have content validity by a panel of experts in nursing research, statistics, and research publication (Selby et al. 1990). Reliability was first assessed for a 1986 pilot test (Selby et al., 1990), for a 1989 study on validity and reliability in selected nursing research journals (Selby-Harrington, Mehta, Jutsum, RiportellaMuller, and Quade 1994), and for the current study. Interrater agreements of 90-92% and 93-97% were reported for the first two studies. Intrarater reliability was established after 12 hours of training conducted by the senior author of the instrument. Intrarater agreement of 88-91% (mean = 90% ± SD 1.1 ) was reported by the authors using seven reviews as the standard; intrarater reliability of 84-96% (mean = 90% ± SD 4.6) was achieved among a group of 11 raters. For this study, test/retest reliability of 98% was established using five reviews. Also, one of every 10 studies was randomly selected to be evaluated by a second reviewer, and 95% intrarater reliability was established. The high inter- and intra-rater agreement tends to ensure that any errors in data collection are random, not biased. Expert consultation in use of the instrument was available throughout the review.

If items were unclear, the policy of the review was to err in the direction of not finding deficiencies. For example, if the purpose of a study was not stated explicitly but could be inferred from information presented anywhere in the article, the study purpose was considered to be adequately described even though a clearer statement would have been preferable. Similarly, if the conceptual framework was unclear but a reference was given, the author was considered to have described it sufficiently. In evaluating the data collection instruments, validity and reliability were judged to be inadequate only if the author made no statement about validity or reliability or explicitly stated that validity and reliability were not assessed.

**RESULTS AND DISCUSSION**

**General Information**

Twenty-seven percent of the articles had one author, 44% had two, 17% had three, and 11% had four. Most (60% of the research was not grant funded or this was not mentioned. Of the 20 funded studies, two had received federal, 15 had received nonfederal, and three had received both federal and nonfederal funding.

**Purpose**

In almost half (45%) of the studies, the stated purpose was to describe an existing situation; in 38% the purpose was to explain existing relationships; in 15% it was to test an intervention; and in one article, the purpose was to conduct a pilot study. Purposes were clearly stated in all but one of the articles.

Twelve of the articles focused on human attitudes toward animals. Nine of the authors were concerned with the motives involved in caring for animals and why human/animal attachments are formed. Four authors were interested in explaining the characteristics of animal providers. Six articles attempted to identify sex-related differences in attitudes.

The effect of animals on health and illness was a particularly popular topic, occurring in 28% of the studies. Four studies involved cardiovascular outcomes; seven, stress outcomes. The effect of animals on human emotions was examined in seven studies, and four authors studied how animals affected behavior. Human/animal interactional processes and relationships were studied in three projects. Animals as a therapeutic intervention were studied by four researchers.

It was difficult to categorize many of the studies. The effect of animal ownership on housing choice or concerns was the subject of five studies. Other topics included the effects of animals on learning, recreation, work, and volunteering.

**Literature Review and Framework**

Literature was cited in the introduction or literature review section of 96% of the articles, and the literature cited clearly justified the need for the research in 87% of the reports. The literature provided unclear or contradictory justification for five of the studies, and five articles reported no justification or no literature. The total number of references listed varied from six to 52 in one study each, with a mean of 22.

Twenty-six percent of the researchers explicitly stated that their studies were based on conceptual or theoretical frameworks, 6% implied a framework, and 68% neither explicitly stated nor implied that their studies were based on any framework or model. Social support theory was the most frequently cited framework (8%). Attachment theory and biopsychosocial theory were used in two studies each. Other models utilized (in one study each) were relaxation theory, classical conditioning, cognitive restructuring, competence, continuity, developmental theory, life events/illness, specialization, Roger's Science of Unitary Human Being, Roy's Adaptation Model, and sex role orientation.

In 11 of the studies, the conceptual frameworks were described sufficiently to be understood. Five authors did not offer adequate descriptions of theoretical models or provide references with sufficient descriptions. The relevance of the frameworks to the studies was clear in 12 of the articles, but in five studies, relevance was not clear.

**Setting and Sampling**

Forty-five percent of the researchers said their samples were taken from urban or city settings; 6% were suburban, and 4%, rural. Six authors named a state only, one cited both rural and suburban, one utilized a sample from outside the United States, and three said their samples came from the entire USA. Three of the articles were unclear about where the sample was taken.

The populations from which samples were selected, or the types of people eligible for inclusion in studies, were clearly described in 48 articles, including demographic characteristics or census.
Sample descriptions were coded in order to establish age groups. Thirteen percent of the studies focused on preschool and school-age children, 17% on young adults, and 32% on the elderly. Two studies covered the span of childhood and 11 covered all adult years. Understudied groups were the middle years (which is consistent with research in most fields), one study; adolescents, two studies; and the institutionalized elderly, three studies.

Forty-four articles presented data from nonprobability and two from probability samples. In six articles, sampling methods were not addressed. Sample sizes ranged from 12 to over 5,000, with 51% below 100 persons. Of the 51 studies in which cooperation of human subjects was needed to get subjects to join the study, the number or percent of refusals or nonresponses was not reported in 32 studies. In one study cooperation of human subjects was not an issue. Only three of the 49 studies involving interventions presented any stated or implied rationale for their sample size and none reported any basis for sample size calculations.

**Research Design**

Research design was stated by authors or implied through description: 43 were nonexperimental, nine were experimental or quasiexperimental. Of the six experimental studies, four met the criteria of manipulation of the independent variable, use of control group, and randomization to groups. One did not describe a control group, two did not randomize, and in one randomization was not clear. The cross-sectional or survey nonexperimental design was the most common data collection method (66%). One study was a retrospective design and three were prospective or cohort designs.

**Data Collection, Validity, and Reliability**

In 48 of the 52 articles, procedures and instruments were described sufficiently to afford an understanding of how the studies were conducted and how at least one major outcome variable was measured. In 45 instances, the definitions, methods, and instruments were judged to be appropriate for measuring all intended variables.

Over half (60%) of the authors stated or implied that at least one major instrument used to collect data had validity; 59% stated or implied reliability; and 20% completely omitted statements about the validity and reliability of instruments.

**Analysis**

Generally, there were few errors committed in reporting analyses of the data: 37 researchers reported a mean for a major independent or dependent variable, and only eight committed the error of not reporting a measure of variability. Raw data were included in all but four studies with small samples, and samples were described statistically in terms of major outcome variables in all but two studies. In studies whose major purpose was to compare groups, descriptive statistics were presented for at least one demographic or other possible confounding variable, and the values for major outcome variables were presented separately for each group in all but two studies. Three quarters of the studies were descriptive. One committed the error of reporting results without enabling the reader to identify the statistical tests used and none neglected to report alpha levels and/or p-values.

**Implications**

Of the 44 nonprobability studies, 15 generalized their findings beyond the sample. Yet 60% of the authors stated the limitations of their studies and 89% made recommendations for further research, conceptual frameworks, or practice; 79% of the implications or recommendations were judged to be appropriately derived from the results of the studies.

**Results**

Qualitative information was gathered on the results of the 52 studies reviewed and those that could fit into loosely constructed categories are summarized here and identified in parentheses. One study used an intervention to elicit a change in attitude toward animals and found that a schoolbased humane education program enhanced animal-related attitudes (Ascione 1992). Most attitude studies described different attitudes or
correlated attitudes with various factors. Descriptive studies found that attitudes changed over the years (Verderber 1991), males and females had different attitudes toward animals (Herzog 1991), most people did not care or were happy about feeding free-ranging cats (Haspel 1990), and owners had an overwhelmingly positive attitude about companion animals (Poresky, Hendrix, Mosier, and Samuelson 1988). Correlational studies found that behavior and attitudes were not significantly related (Miller and Lago 1990), attitudes were affected by social class and spousal attitude (Gage 1988), parental attitudes influenced children's attitudes (Kidd and Kidd 1990), and the ability to own an animal did not affect attitude toward animals (Mihalski, Jones, and Maxwell 1988).

Studies describing motives for caring for animals and attachments to animals found that the primary reason for feeding birds was aesthetics (Horvath and Roelans 1991). Motives for living with birds were social and cognitive (Loughlin and Dowrick 1993), and people fed free-ranging cats because they felt they were hungry (Haspel and Calhoon 1990). Other researchers found that urban children were more attached to animals than rural children were, and there was no significant difference between males and females in attachment to animals (Stevens 1990). Owners were most attached to their first companion animals (Nielson and Delude 1989), and owners had a lifetime of consistent ownership (Nielson and Delude 1989). Further studies found that a child's age and maternal employment were related to the child's involvement with animals (Melson 1988); involvement with animals was positively associated with other nonschool activities (Melson 1988), and gender did not affect motivation to have an animal companion (Siegmund and Biermann 1988).

Characteristics of animal owners were also studied. Researchers found that 63% of households included pets (Marx, Stallones, Garrity, and Johnson 1988). Animal ownership increased with the child's age (Melson 1988), and children with animals had more activities and interests than children who did not have animals (Kidd and Kidd 1990). whether or not a child had an animal was a consequence of many different variables (Paul and Serpell 1992, Melson 1988). One study found that girls talked to animals more than boys (Nielson and Delude 1989), and one revealed the negative aspects of childhood animal ownership (Bryant 1990).

Four studies involving elders found positive outcomes of human/companion animal interactions (Calvert 1989; Mahalski, Jones, and Maxwell 1988; Netting, Wilson and Frugue 1988; Peretti 1990). One study, however, revealed the difficulties of animal ownership for the elderly (Verderber 1991).

Of the studies examining gender differences in relation to animal attachment, attrition, choice, or motivation, most did not find any significant differences (Lago 1989, Siegmund and Biermann 1988, Stevens 1990). One researcher, however, did find a difference between male and female attitudes toward animals (Herzog, Betchart, and Pittman 1991).

Of the eleven studies dealing with animal effects on health and illness, the majority found a positive association between animal contact and health (Allen, Blascovich, Tomaka, and Kelsey 1991; DeShriver and Riddick 1990; Friedmann, Locker, and Lockwood 1993; Garrity, Stallones, Marx, and Johnson 1989; Lago, Delaney, Miller, and Grill 1989; Wilson 1991). Three did not find such an association, however (Grossberg, Alf, and Vormbrock 1988; Melody and Lago 1990; Stallones, Marx, Garrity, and Johnson 1990). Of those dealing specifically with stress effects, three found that human/animal interactions decreased stress levels (Allen, Blascovich, Tomaka, and Kelsey 1991; DeShriver and Riddick 1990; Wilson 1991), and one found no significant difference between owners and nonowners (Watson and Weinstein 1993). The studies involving animals as therapeutic interventions were unanimous in reporting successes (Calvert 1989; Holcomb and Meacham 1989; Kongable, Buckwalter, and Stolley 1989).

Researchers also found that animal interactions tended to make people happier, more relaxed, more secure, more affectionate, more alert, and less lonely (Bryant 1990, Calvert 1989, Rossbach and Wilson 1992). One article, however, suggested that negative emotions such as grief and distress can also be associated with companion animals (Bryant 1990). Animals were found to increase touch, talk, and smiling (Nielson and Delude 1989). Rogers, Hart, and Boltz (1993) found that dog owners exercise more and were more satisfied
with their social, physical, and emotional states. Two researchers reported the successful relationships that owners had with animals, often viewing them as members of the family (Barker and Barker 1988, Turner and Stammbach-Geerring 1990).

**Limitations of This Review**

The review included 52 reports representing six years of research in the area of human/companion animal relationships. The articles included were not a random sample but were convenient to the researcher. The review may have unintentionally omitted relevant reports. For example, a number of studies have been published in French and German journals, but only English language research reports were included here. Also, theoretical works were excluded because the study was intended to evaluate only research reports. Finally, some studies were evaluated as formal research reports when that may not have been the intent of the authors. To counteract the possibility of being overcritical and to offset overestimation of methodological deficiencies, the reviewer consistently erred on the side of not identifying deficiencies; this may have resulted in an underestimation of methodological problems.

**SUMMARY AND RECOMMENDATIONS**

This study evaluated the quality of 52 articles on human/companion animal research published over the period 1988 to 1993. This review, while not comprehensive, points out the significant gains made in knowledge of the impact of the human relationship with companion animals. The review primarily examined the rigor of the research included. The research consisted primarily of nonexperimental studies. Although samples were large enough, most were nonprobability, nongeneralizable samples. However, the use of nonprobability sampling is not itself problematic, the authors of more than 25% of those studies generalized their results inappropriately, beyond the sample.

A number of other problems were identified. Future studies should use sample sizes based on mathematical calculations and power analysis, use theoretical frameworks, and assess the reliability and validity of measurement instruments. This review provides a beginning evaluation of the research on human/companion animal relationships that is likely to influence practice. This type of review is time consuming and, unfortunately, not readily fundable. Nevertheless, it is necessary for researchers and practitioners in the human/companion animal field to evaluate the research, take actions to improve it, and assess the results of these actions. Ultimately, this will result in improved practice.

**REFERENCES**


APPENDIX: SELBY RESEARCH ASSESSMENT FORM III (RAF III)

Revised January 1993

Maija L. Selby, Dr. RH., R.N., C.P.N.P.

School of Nursing

University of North Carolina at Greensboro

Greensboro, North Carolina 27412

Phone: (919) 334-5010

Copyright 1993 Maija L. Selby

*This document may not be copied without permission of the author.*

**ABSTRACTOR:** Attach photocopied article. Use highlighter and make notes on article as needed. For RAF items requiring explanation or questions, on the RAF please unite PAGE number, COLUMN number, and PARAGRAPH number M article so that relevant information can be found easily (count top of page as Paragraph One even if it starts in the middle of a paragraph).

*For all items in which you do not have sufficient evidence to cite a "wrongdoing," give the authors of the article the "benefit of the doubt." However, if you are totally unsure of an item, refer questions to the PI. DO NOT GUESS.*
SECTION I:
GENERAL INFORMATION

1. Abstractor Name (see code list)

2. Article i.d.

3. Journal (see code list)

4. Year

5. Number of authors

6. Highest degree of first (or only) author

7. Is there a doctorally prepared (PhD, DrPH, MD, DO, LLD, DNS, DNSc, DSN, ScD, etc.) author after the first author?

8. Nursing/Medical of first (or only) author?

9. Is there an RN (RN, ADN, BSN, ND, DNS, DNSc, NP, FNP, ANP, PNP, etc.) author after the first author?

10. Is there an MD/DO author after the first author?

11. Grant funding acknowledged? ABSTRACTOR: usually at bottom of first page with author listing, or at end in Acknowledgments section.

1. federal only
2. both federal and non-federal
3. none listed

SECTION II:
PURPOSE, LITERATURE REVIEW, FRAMEWORK

1. Is purpose statement clear? ABSTRACTOR: list purpose in item la below.

1. explicitly stated in article or abstract
2. not explicitly stated but implied
3. totally unclear or absent

2. Stated purpose is:

1. to conduct a pilot test (instrument, methods, etc.)
2. to describe existing situation
3. to explain or test differences or relationships in existing situation
4. to test an intervention
5. other; describe
6. totally unclear or absent

3. Is ANY literature cited in INTRODUCTION, LITERATURE REVIEW, or ANYWHERE BEFORE METHODS SECTION?

4. Does cited literature justify need for this study?

5. Total number of references in article (count reference list):

6. Is study based on conceptual or theoretical framework or model?

1. yes, explicitly states that study based on specific framework or model
2. implies that study based on specific framework or model
3. does not explicitly state nor imply that study based on specific framework or model

7. Identify conceptual or theoretical framework (to be coded later)

8. Is conceptual or theoretical framework described sufficiently for you to understand it?

9. Is ANY attempt (good or bad) made to explain the relevance of framework to this study?
10. Is framework relevant to this study? 1. yes, framework clearly is relevant, based on information provided in article  
2. unable to determine relevance based on information provided in article  
3. framework clearly is not relevant or is contradictory to this study, based on information in article  
4. N/A; no framework stated or implied

SECTION III: SETTING & SAMPLING
1. Is the setting for the study mentioned or implied?  
2. Specific setting (where they say their sample is from). List whether RURAL, URBAN, SUBURBAN; STATE or CITY; whatever information you can gather. (To be coded later)  
3. Is sampling frame or population from which sample was selected mentioned (at all, anything any description of types of people eligible for inclusion in study)?

4. Are ANY demographic characteristics of sample or census described (in methods or in results section)?

For sampling questions, if study is multi-part with more than 1 sample for different research purposes, give the authors credit for the BEST that they did. Describe any unique aspects in 4a below.

4a. Describe sample (will not be coded)

5. Was sampling method discussed or described at all?

6. Sampling method that actually appears to have been used  
1. probability (Do not confuse random sampling with random assignment)  
2. nonprobability  
3. census (entire population)  
4. unable to tell

7. Total completed sample size (sample #1) 888888. unclear or conflicting information; DESCRIBE 999999. no information on sample size

8. Total completed sample size (sample #2) 888888. unclear or conflicting information; DESCRIBE 999999. no information on sample size

9. In studies in which cooperation of human subjects was needed to get subjects to join the study, was the number or percent of refusals or nonresponses given?

10. If INTERVENTION, does the article give the number of withdrawals or cases lost after the study began, and the reason(s for withdrawals or cases lost

11. What rationale was provided for planning sample size?

12. Were sample size calculations reported to be based on  
1. descriptive statistics?  
2. inferential statistical tests (e.g., differences between groups)?  
3. both descriptive and inferential statistics?  
4. no mathematical calculations mentioned 8. N/A; no sampling

13. For studies whose main purpose is to describe differences between groups (inferential statistics or statistical tests), how was power and effect size reported?  
1. Power only reported  
2. Effect size only reported  
3. Power and effect size reported
4. No power reported and no effect size reported. 8. N/A; purpose of study was to describe, not test; or, no sampling

SECTION IV: RESEARCH DESIGN
1. Research design used, as stated by authors or if not stated, implied through description
   1. experiment
   2. quasiexperiment or pre-experiment
   3. nonexperiment
   4. not stated or implied
   5. other; describe

2. Was there manipulation of the independent variable (intervention)?
   3. Did they use a control or comparison group?
   4. Was there randomization to groups? ABSTRACTOR: Do not confuse with random sampling.

5. Type of experiment or quasiexperiment
   1. 1 group
   2. 2 groups
   3. 3 or more groups
   4. other; describe
   8. N/A; nonexperimental

6. Type of nonexperimental design
   1. retrospective (past-oriented)
   2. cross-sectional or survey (present-oriented, at time of data collection)
   3. prospective or cohort (future oriented, follows forward)
   4. used retrospective data to follow prospectively
   6. other, specify
   8. not described clearly enough to categorize
   88. not described at all

88. N/A; experiment or quasiexperiment

SECTION V: DATA COLLECTION, VALIDITY, RELIABILITY
1. Were procedures described sufficiently to understand how study was conducted (e.g., surveys, interviews, secondary analysis, etc.)

2. Are operational definitions or descriptions of methods or instruments sufficiently clear to understand how the researchers measured the MAJOR outcome variable(s)? ABSTRACTOR: Do not evaluate quality of definition, or instrument. A description may be clear, but clearly wrong. Evaluate quality in item #3.
   1. Yes, at least fairly clear for all major variables.
   2. At least fairly clear for at least one major variable, but unclear or missing for other(s). Describe:
   3. Unclear or missing for all major variables. Describe:

3. Are definitions, methods, or instruments appropriate for measuring the MAJOR outcome variable(s)?
4. Did author(s) state or imply that data collection instruments for major variables have validity (any kind)?

5. Did author(s) state or imply that data collection instruments for major variables have reliability (any kind)?

SECTION VI: ANALYSIS: Part A

1. If the mean is reported for a major independent or dependent variable, does the author COMMIT THE ERROR of NOT REPORTING a measure of variability (standard deviation, standard error, variance, or range)?

2. If proportions or percents are reported in a very small sample or subsample (N < 20), does author COMMIT THE ERROR of NOT REPORTING information to enable you to see raw frequencies?

3. Is sample or census described STATISTICALLY in terms of major outcome variable from problem statement, research question(s), and/or hypothesis(es)? ABSTRACTOR: article must report in text or table a measure of central tendency (unless nominal) or variability for major variable.

4. If comparison of groups is a major purpose, are DESCRIPTIVE STATISTICS presented for at least one demographic or other possible confounding variable, SEPARATELY FOR EACH GROUP?

5. If comparison of groups is a major purpose, is VALUE for major outcome variable presented SEPARATELY FOR EACH GROUP?

6. Categorize the purpose of the study: I. descriptive purpose only
   2. descriptive with incidental statistical testing
   3. statistical testing is a major purpose
   4. totally unclear purpose

7. How many statistical tests are reported or implied by a statement such as (p < .05), "were statistically significant," etc.?

8. Does author COMMIT THE ERROR of reporting results of statistical testing for major purpose of study without enabling reader to identify statistical test used (either in methods or results)? e.g., "the results were not statistically significant" or the hypothesis was rejected" or "the differences were significant" but you can not determine what test was used.

9. How does author handle reporting of alpha (level of significance) and p-values for MAJOR statistical test?
   1. reports alpha level and p-values, or p-values only (p = exact number/-WITHOUT reporting alpha IS OK; p = ns WITHOUT reporting alpha is NOT OK)
   2. COMMITS ERROR of NOT doing it as described as "OK" in option I above. N/A; no statistical tests reported

10. List major findings (should be in relation to purpose of study):

SECTION VII: IMPLICATIONS

1. If NO statistical significance is found with a small sample or subgroup, does author place considerable confidence in the lack of significance? ABSTRACTOR: e.g., does the author make a "big deal" about there being no difference or no relationship, and not refer to the fact that the small sample may have caused it?

2. Were results generalized beyond the sampling frame or census?

3. Were results generalized beyond sample? (if census rather than a sample was used, consider the census to be the sample for the study)

4. Were ANY limitations of study stated?

5. List serious limitations, if any (stated or unstated)(to be coded)

6. Are there ANY stated implications or recommendations for research, theory, practice, education, policy, etc.? ABSTRACTOR: Do the authors say ANYTHING about what to do because of the findings-e.g., change our base of knowledge because they found the "right" answer, change practice or policy, do more research?

7. Are ANY implications or recommendations stated for future research?
8. Are ANY implications or recommendations stated for theoretical framework?

9. Are ANY implications or recommendations stated for practice, education, or policy?

10. Are the implications or recommendations as a whole derived from the results of THIS study? Do not count overgeneralization here; just look for blatant contradictions between study results and what the authors say should be done, or for a set of implications and recommendations that could have been made without doing the study.

SECTION VIII: ANALYSIS, Part B (specific statistical tests and power evaluations) ABSTRACTOR: Evaluate whether a reported statistical test is used appropriately or inappropriately, based on the assumptions underlying the use of each test. Use the Selby Chart, LeGault

Instructions, Daniel or Remington and Schork textbooks, and/or your statistical consultant to make these evaluations. Do NOT count violations of assumption of normality or homoscedasticity as a violation. Please state how each of the following tests are used: confidence interval, Z-test (any kind), McNemar Test, Fisher Exact Test, Binomial Test, Mann-Whitney, analysis of variance (ANOVA or F test), analysis of covariance (ANCOVA), Pearson Product Moment (Pearson r, bivariate regression), Spearman Rho, Kendall Tau, Median Test, Sign Test, Wilcoxon Test, Cochran Q, Multiple Regression, Factor Analysis, Kolmogorov-Smirnov Test, Kruskal-Wallis Test, and t-test (any kind).

1. If Chi Square is used, is it used incorrectly with related (hot independent) samples?

2. Is t-test for related samples used when independent sample t-test required? ABSTRACTOR: assume that independent sample t-test

is used unless otherwise stated

3. Is t-test for independent samples used when related sample t-test required? ABSTRACTOR: assume that independent sample t-test is used unless otherwise stated

4. Are multiple t-tests used (e.g., Group A vs. Group B, Group A vs. Group C, Group B vs. Group C?) without stating that alpha was reduced?

5. Are repeated observations analyzed as independent (any test)? ABSTRACTOR: e.g., WHEN N =6 INFANTS BUT 300 CRYING EPISODES ARE ANALYZED OR N=10 PEOPLE MEASURED BEFORE AND AFTER AND N=20 SCORES ARE ANALYZED.

Items 28-31 are needed for the first major hypothesis for finding power from Cohen tables. Write in the value AND fill in the code about whether information is provided.

Alpha value (use .05 unless otherwise stated in article):
alpha explicitly stated
alpha not stated but p-values are provided 3. no alpha and no p-values; therefore alpha assumed to be .05
no stat tests used
Name of test:
name of test stated
name of test not stated 8. no stat tests used
Value of test:
value of test stated
value of test not stated 8. no stat tests used
Sample size:
sample size stated
sample size not stated 8. no stat tests used
Power for SMALL effect
Power for MEDIUM effect
Power for LARGE effect

Does this study FAIL TO REJECT the major null hypothesis for which you found power, above?
ABSTRACTOR: YOU HAVE REACHED THE END. THANK YOU FOR YOUR EFFORT!