8-1974

Polygraph Use in Increasing the Reliability of Self-Recorded Data

Stephen M. Zweibach
Western Michigan University

Recommended Citation
https://scholarworks.wmich.edu/masters_theses/2624
POLYGRAPH USE IN INCREASING THE
RELIABILITY OF SELF-RECORDED DATA

by

Stephen M. Zweibach

A Thesis
Submitted to the
Faculty of The Graduate College
in partial fulfillment
of the
Degree of Master of Arts

Western Michigan University
Kalamazoo, Michigan
August 1974
ACKNOWLEDGEMENTS

In preparing this thesis, I have benefited greatly from the encouragement, advice and constructive criticism of Dr. Paul T. Mountjoy, Dr. Hermann A. Peine and Dr. Chris Koronakos. In addition, I would like to acknowledge the support and assistance of Dr. Neil D. Kent in the early stages of implementing this study. My thanks go to them as well as to the many others in the Psychology Department at Western Michigan University who have contributed to my intellectual training, knowledge and understanding of behavior analysis applied in the natural environment.

Stephen M. Zweibach
INFORMATION TO USERS

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

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

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

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

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

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from “photographs” if essential to the understanding of the dissertation. Silver prints of “photographs” may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
MASTERS THESIS

ZWEIBACH, Stephen M.
POLYGRAPH USE IN INCREASING THE RELIABILITY OF 
SELF-RECORDED DATA.

Western Michigan University, M.A., 1974
Psychology, clinical

University Microfilms, A XEROX Company, Ann Arbor, Michigan
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>1</td>
</tr>
<tr>
<td>II</td>
<td>5</td>
</tr>
<tr>
<td>Subjects</td>
<td>5</td>
</tr>
<tr>
<td>Setting</td>
<td>5</td>
</tr>
<tr>
<td>Polygraph Examinations</td>
<td>6</td>
</tr>
<tr>
<td>Target Behaviors</td>
<td>7</td>
</tr>
<tr>
<td>Procedure</td>
<td>8</td>
</tr>
<tr>
<td>Baseline 1</td>
<td>9</td>
</tr>
<tr>
<td>Self-recording 1</td>
<td>9</td>
</tr>
<tr>
<td>Polygraph sessions</td>
<td>10</td>
</tr>
<tr>
<td>Self-recording 2 and baseline 2</td>
<td>10</td>
</tr>
<tr>
<td>Independent Observations</td>
<td>11</td>
</tr>
<tr>
<td>III</td>
<td>13</td>
</tr>
<tr>
<td>Reactivity</td>
<td>13</td>
</tr>
<tr>
<td>Reliability</td>
<td>16</td>
</tr>
<tr>
<td>IV</td>
<td>23</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>23</td>
</tr>
<tr>
<td>V</td>
<td>26</td>
</tr>
<tr>
<td>APPENDIX A</td>
<td>26</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
INTRODUCTION

In behavior modification research and intervention direct observations of behavior are utilized to assess behavior change. Reliable data are essential for any result to be conclusive. When the behavior in question can be observed by additional independent observers, as is often the case, data on reliability are relatively easy to obtain. Many behaviors, however, can be observed and recorded only by the client. In these cases, where independent observations of overt behavior are difficult or impossible, the collection of data on the reliability of the subject's self-recording is precluded. Indeed, many early studies employing self-recording did not even report reliability (Cautela, 1966; Fox, 1962; Goldiamond, 1965; Stuart, 1967).

Throughout the literature, various authors have noted that the reliability of self-recording has not been adequately demonstrated; among these are Kanfer (1970), Nelson & McReynolds (1971), Simkins (1971), and Tighe & Elliott (1968). Research has done little to counter this impression. When unobtrusive independent observers have been employed, in addition to the subjects monitoring their own behavior, the subjects' recordings have invariably been found to be unreliable (Broden, Hall & Mitts, 1971; Fixsen, Phillips & Wolf, 1972; Herbert & Baer, 1972; Lipinski & Nelson, 1974; McFall, 1970; Risley & Hart, 1968).

In spite of the difficulties in determining reliability and
often shown unreliability, self-recording has been and continues to
be used, particularly in therapeutic situations. Basically, there
are two reasons for this. First, there is often simply no other
measure available, short of having an observer follow a subject
twenty-four hours daily. Thus, for example, self-recording is
almost always found in studies dealing with smoking (Axelrod, Hall,
Weis & Rohrer, 1974; Chapman, Smith & Layden, 1971; McFall & Hammen,
1972; Rutner, 1967) and weight loss (Stollak, 1967; Stuart, 1971).
Second, many studies appear to have capitalized on reactive effects
of self-recording on the operant level of a self-recorded behavior.
That is to say that self-recording is used as an intervention
strategy which can effectively alter behavior (Tharp & Wetzel, 1969).
Reactive effects, however, have not always been observed (Hall, 1972;
Mahoney, Moura & Wade, 1973; Stollak, 1967), and in some cases have
been reported to attenuate with time (Broden, Hall & Mitts, 1971;
Fixsen, Phillips & Wolf, 1972; Stuart, 1971). Nevertheless, as
Lipinski & Nelson (1974) note, the direction of behavior change
produced by self-recording is often in a therapeutic direction
(Johnson & White, 1971; McFall, 1970; McFall & Hammen, 1972).
Further, it appears that reliably observing one's own behavior is
not essential for changing that behavior, as pointed out by Broden,
Hall & Mitts (1971) and Herbert & Baer (1972).

Although self-recording plays an integral role in behavioral
change, the behavior change evidenced in self-report measures
cannot necessarily be considered to reflect an actual change in
non-verbal behavior in the absence of corroborative data (Kazdin,
1974). Thus, in the usage of self-recorded data for assessment purposes, one must attempt to increase or at least monitor the reliability of the data. Several investigators have attempted to monitor and maintain reliable self-report data. Powell & Azrin (1968), employing their participant-observer technique, demonstrated 88% and 93% agreement of self-recorded smoking behavior for two subjects. Ober (1968) reported a correlation of .94 of self-recorded smoking with reports of two "friends" familiar with the subject's smoking habits, however the observers' recording methods were not elaborated. Finally, Gerson & Lanyon (1972) attempted to insure reliability with a novel approach. They screened their subjects for dishonesty in reporting facts about themselves with the K and L scales of the Minnesota Multiphasic Personality Inventory. Unfortunately, data needed to validate the usage of these scales was not presented.

It is clear that methods for increasing the reliability of self-recorded data are needed, particularly for those situations where independent observations are not possible and self-recorded data provide the only dependent variable for assessment. In criminal and civil investigations, the lie-detector or polygraph is frequently employed to increase the reliability of self-reported information. The Polygraph technique is reported to be a very reliable aid in the determination of truth or deception (Reid & Inbau, 1966). In view of common acceptance of this dogma, it would seem that the administration of a polygraph examination questioning
the diligence of a subject in self-recording would serve to increase 
the reliability of that recording.

The present study had two main objectives. First, an attempt 
was made to increase the reliability of self-recording through the 
administration of a polygraph examination. Second, the reactive 
nature of self-recording was examined by observing the effects on 
the operant level of the self-recorded behavior, as well as the 
duration of the effects with time.
METHOD

Subjects

Three college students taking advanced undergraduate psychology classes at Western Michigan University were asked to participate in an experiment dealing with television viewing and which entailed behavioral recording. There were two males and one female (S2). In return for their participation in the present study they received academic credit.

Setting

Sessions took place in a psychology laboratory equipped with a Zenith 24 inch television set located on a shelf 30 inches from the floor. A chair with a writing arm was fixed to the floor seven feet before the television. Each subject was provided with a pencil and a form for interval recording of behaviors. A Panasonic cassette tape recorder, model number RQ-309S, located in an adjacent equipment area, played a pre-recorded 2900 Hz tone through the television speaker every six seconds during the session. The tape began with the message, "At the tone the first interval will begin," immediately followed by the first tone. Thereafter, the number of every tenth interval was announced immediately following the tone signalling the start of that interval. These numbers corresponded to numbered spaces on the recording forms. During baseline and return-to-baseline sessions, a one page form containing three hundred numbered...
spaces was used. In all other sessions, a two page form, providing two spaces per number, was used.

A Sony Videocorder AV-3400, located in the adjacent equipment area, was used to videotape the subject during each session. The camera was concealed behind bookshelves holding the television. A microphone was concealed within the television, beside its speaker. Thus, videotapes included a picture of the subject, the sound of the television game show and the interval tones played during the session. In addition, a Sony cassette-corder, model TC-66, was placed on a table beside the television.

Polygraph Examinations

Polygraph examinations were administered with a Model 6303 Keeler Polygraph, located behind a desk in an adjacent room. A chair was provided in front of the desk, with its back to it, for the subjects of the examinations. Each subject was fitted with a cardio cuff, respiration device and finger electrodes, attached to the polygraph. Sitting in the chair, with the experimenter sitting at the desk, each subject was asked to answer the following questions with either a yes or no reply:

1. Is your name ________? (insert subject's name)
2. Do you attend Western Michigan University?
3. Do you intend to lie to me on questions regarding your recording of face-touching today?
4. Were you alert while watching the television today?
5. Were you daydreaming at any time?
6. Did you attend carefully to your face-touching behavior at all times?
7. Did you record every instance of face-touching that occurred, as defined by the definition?
8. Are you sure of that?
9. Were you attending to your face-touching during commercials?
10. Were you at any time so absorbed in the show that you may have failed to record some face-touching?
11. Do you ever listen to the radio?
12. Have you lied in answering any questions on this test?
13. Have you done anything to try to impair the value of this test?

There was a fifteen second pause between every reply and the next question. While asking the questions, the experimenter adjusted the polygraph, made notes, and labeled the polygraph chart record.

Target Behaviors

Subjects were asked to record audience participation on a television game show, defined as any hand-clapping, laughing, hooting, talking, or other verbalizations not produced by the participants of the show or speakers on commercial announcements. Applause was recorded in an interval regardless of whether it began in a
preceding interval or not.

The self-recorded behavior, face-touching, was defined as touching the face, head, hair or neck with a hand, pen, pencil or paper. Objects were considered equivalent to face or hand, for example, touching a cigarette to mouth. If any instance of face-touching was evident within an interval, it was counted. Thus, if one rested one's chin on one's hand for ten intervals, all of them would be scored affirmatively for face-touching.

Procedure

During the first session, subjects were given a questionnaire, unrelated to the experiment, dealing with their television viewing habits (see Appendix A). This was done to maintain the appearance of a study dealing with television viewing. They were then instructed as follows:

"This study will require you to watch a television game show at the same time every day for a half hour, recording certain events. A tone will sound every six seconds, signalling the start of an interval. We would like you to record the presence or absence of audience participation in these intervals. If audience participation occurs at any time during the interval, put a letter A in the box which corresponds to that interval. If not, place a dash in that box when the next tone sounds."

The definition of audience participation was presented and explained. The remainder of the session was a training session, providing practice in recording audience participation. The experimenter remained present, providing immediate feedback on appropriate and inappropriate recording on the part of the subject. No further training or
feedback was provided, with the exception of S1, who was advised of poor reliability between his recording and that of the experimenter prior to the second baseline session.

At the start of all sessions, when the subject was seated the experimenter turned on the television. He next entered the equipment area, turned on the tone-tape and videocorder, and emerged. Finally, he set the recorder beside the television to record, and then left the room for the remainder of the session. No other persons were present in the room during the sessions.

The experiment was a multiple baseline design. S2 began self-recording sessions one day after S1 began self-recording sessions. S3 began self-recording sessions two days after S1. Likewise, subsequent condition changes were staggered for the three subjects.

**Baseline 1**

During the first four sessions for subjects S1 and S2, and five sessions for subject S3, the subjects were instructed to watch the thirty minute program, recording audience participation as accurately as possible. They were informed that reliability would be assessed, from then on, by means of a tape-recorded soundtrack of the session, made with the visible cassette recorder.

**Self-recording 1**

At the start of the fifth session, subjects were informed that the study would require them to accurately record face-touching
behaviors in the same intervals as and simultaneous with the recording of audience participation. They were provided with the two page recording form and instructed to place a letter F in any interval in which face-touching occurred, and a dash in those in which face-touching did not occur before the next tone. The remainder of the session was a training session on self-recording face-touching along with audience participation. The target behavior was described and the subjects practiced recording both behaviors at once, in the presence of, and with feedback from, the experimenter. For the next four sessions the subjects were instructed to watch the program, recording both behaviors as accurately as possible.

Polygraph sessions

At the start of the next session, the subjects were informed that they would take a short lie-detector test on how accurately they recorded and attended to their face-touching behavior in that session, so as to be sure that we were getting accurate and reliable data. Following this and each of the next two sessions, each subject was led into the adjacent room and given a polygraph examination, as described above. The subjects were advised that the polygraph chart record would be analyzed later by an expert. No feedback as to the actual polygraph responses was given at any time. In fact, the data was not used at all.

Self-recording 2 and baseline 2

Prior to the start of the next session, subjects were informed
that the polygraph results showed that they had been recording acceptably and that therefore the test would not be used any longer. They were again instructed to watch television and record the two behaviors, maintaining their level of accuracy, for the next four sessions. Finally, in the remaining four sessions, subjects were again given the one-page recording forms and instructed to record audience participation only.

Following the last session, the subjects were asked to fill out the following questionnaire:

1. What aspects (if any) made you suspicious of the experiment?
2. To what extent do you feel polygraph examinations can detect your truthfulness?
3. Have you formed any ideas as to the actual purpose of the experiment? If so, what are they and when did you form them?

Independent Observations

Following each session, the videotape was observed by the experimenter. Guided by the tones recorded on the videotape, both audience participation and face-touching of the subject were recorded, using the two-page recording form. Two additional observers were trained by the experimenter, in the same manner as the subjects, to a criterion level of 90 percent agreement or above, for two consecutive sessions. One tape, randomly chosen from each subject's tapes under a condition, was observed by an additional observer. The
observer was unaware of which condition was being viewed.

The first and last five minutes of each session were not in­
cluded for data analysis throughout the experiment, to control for
possible warm-up or fatigue effects, respectively. The central
twenty minutes were treated as two ten-minute blocks of one hundred
six-second intervals, each block providing one data point. In this
way, any consistent variability within sessions could be assessed.
Reliability between the experimenter and the independent observer for
a given tape was assessed, based upon the ratio of the number of
agreements over the total number of agreements and disagreements. In
no case was the percentage of agreement less than 92 percent for
audience participation or 98 percent for face-touching behavior.
RESULTS AND DISCUSSION

On the final questionnaire all subjects reported some degree of suspicion as to the nature of the experiment initially, although none of them formed any ideas concerning the actual purpose of the experiment. In addition, they reported a strong belief in the utility of polygraph examinations in detecting truthfulness. Finally, all of the subjects reported being unaware of the presence of the hidden camera. These latter reports were confirmed to some extent by the experimenter's observations of "surprise" on the part of the subjects upon revealing the camera to them after the last session.

Reactivity

The frequencies of face-touching observed by the experimenter are plotted in Figure 1. Mean frequencies of face-touching showed a 61 percent decrease for S1, a 51 percent decrease for S2 and a 57 percent decrease for S3, during self-recording 1 sessions, as compared to baseline 1 sessions. These results suggest that self-recording is indeed a reactive procedure, thus corroborating such findings by Broden, Hall & Mitts (1971), Gottman & McFall (1972), Johnson & White (1971), McFall (1970), and McFall & Hammen (1972).

The reactive effect of self-recording was clearly temporary. The frequency of face-touching of S1 and S2 increased with the termination of self-recording, surpassing baseline 1 levels. S3 showed an increase in face-touching when self-recording ceased, to a
Figure 1 - Frequency of Face-Touching
level substantially greater than that during the preceding self-recording condition, but nevertheless lower than baseline 1 performance. In general, these results show a temporary effectiveness of self-recording, consistent with that seen in Broden, Hall & Mitt's (1971) study with studying and talking. They are incongruent, however, with the findings of McFall (1970) and McFall & Hammenn (1972) who observed maintenance of some reductions in smoking following termination of self-recording, as well as those of Gottman & McFall (1972) who found that an increased rate of talking was maintained when self-recording ceased. The failure of S3 to attain baseline 1 levels of face-touching, however, does suggest some permanence of the effect, consistent with the latter findings. A longer baseline 2 condition might have clarified this point, however, time limitations precluded this in the present study.

A slight further attenuation of face-touching was apparent during polygraph sessions, as seen in Figure 1. This appears to be very similar to a further decrement in face-touching observed by Lipinski & Nelson (1974) when their subjects were aware that the reliability of their face-touching was being assessed. This suggests that the polygraph may function analagously to overt reliability assessment in facilitating the reactive effect of self-recording.

The frequency of face-touching of S2 gradually increased to baseline 1 levels during the second self-recording condition. This suggests that the reactive nature of self-recording might attenuate with time, as observed by Stuart (1971) and Fixsen, Phillips & Wolf (1972). Unfortunately, only slight increments are seen in the face-
touching of S1 and S3 during self-recording 2. Again, perhaps more sessions of self-recording would have brought about attenuation of the reactive effect for S1 and S3.

In general, the direction of the changes produced by self-recording differs, perhaps as a function of the instructions given to the self-observer or the value judgement attached to the behavior under observation (Kanfer, 1970). Thus, McFall (1970) and Gottman & McFall (1972) differentially produced changes in smoking and class participation, respectively, by the nature and context of their recording instructions to the subjects. Broden, Hall & Mits (1971), with self-recording, showed increases in an appropriate behavior, studying, while showing decreases in an inappropriate behavior, talking out. Likewise, Johnson & White (1971) increased study behavior. In the present study, subjects were not given any differential instructions, nor were explicit value judgements placed on the target behavior. It is possible, however, that internalized value judgements might have placed face-touching in a negative category for the subjects. Hence, as Lipinski & Nelson (1974) suggest, decreasing face-touching might be self-reinforcing.

Reliability

Figure 2 represents the percentages of agreement between each subject's recordings of audience participation and those of the experimenter. The reliability values were obtained by dividing the number of agreements by the total number of agreements plus disagreements and multiplying the quotient by 100. Since audience partici-
Figure 2 - Percent Agreement of Audience Participation Recording

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
pation generally occurred in forty to sixty percent of the intervals, these values are not subject to artificial inflation by extreme rates of the behavior (Hawkins & Dotson, 1972). Further, the probability of obtaining high reliability by chance is very low. The recording of all three subjects was consistently highly reliable, nearly always greater than 90% agreement, excepting the first two baseline sessions of SI, when, as mentioned above, feedback was presented to develop high reliability. This finding is not surprising in view of the fact that these observations were made under continuous overt reliability assessment, which has been shown to maintain reliable observational recording (Reid, 1970).

As indicated in Figure 1, the frequency of face-touching decreased to very low rates during self-recording. Bijou, Peterson & Ault (1968) point out that when behavior is at a very low rate the observers could disagree on the occurrence of the behavior yet still show high reliability due to their agreement on the large number of intervals where no behavior was observed. Accordingly, as Hawkins & Dotson (1972) suggest, scored-interval agreement was used for assessing the reliability of face-touching self-recording. In calculating scored-interval agreement, all intervals in which neither observer scored the behavior as occurring are ignored.

Figure 3 presents the percentages of agreement of scored intervals for the recording of face-touching. A sizable amount of variability exists in the reliability of self-recording depicted here. One should bear in mind that not only is scored-interval agreement a very stringent test of observer reliability, but also
Figure 3 - Reliability of Self-recording
that scores become highly variable at extremely low frequencies of behavior (Hawkins & Dotson, 1972). There is, however, consistent variability found in the self-recording 1 condition of S3, with reliability scores lower in the second block of each session. Administration of the polygraph examinations removed within session variability from the self-recording of this and the other subjects. Thus, the polygraph procedure appears to be useful in maintaining consistent self-recording within a session.

S1, S2, and S3 had mean scored-interval agreement scores of 50.9%, 74.1% and 71.6%, respectively, during self-recording 1 sessions. Reliability increased to 100% during polygraph sessions for all subjects, with the exception of S3, who recorded with 85% reliability during the first session. Reliability decreased following termination of the polygraph sessions, with S1, S2 and S3 providing mean scored-interval agreement scores of 40%, 74.3% and 68.9%, respectively, during self-recording 2 sessions. One should note that face-touching did not occur at all in five blocks for S1 and in one block for S2, precluding the calculation of scored-interval agreement scores for these intervals. It should also be noted that the unreliability evidenced during the self-recording conditions arose either as a result of a subject's failure to record a behavior or a subject's recording in a wrong, usually adjacent, interval. As a result, scored-interval reliability could be low while the frequency recorded per block by the subject might agree with that observed by the experimenter. Reliability values calculated by dividing the smaller frequency reported per block are also presented in Figure 3.
These data suggest that roughly half of the unreliability apparent in the scored-interval agreement data was due to inaccurate interval recording, rather than due to a failure to record behaviors. Nevertheless, the effect of the polygraph is clear.

The results tend to corroborate the findings of other investigators in that self-recording was found to be unreliable. In particular, they serve to replicate Lipinski & Nelson's (1974) finding of unreliable self-recording of face-touching. The multiple baseline employed in the present study appears to be an improvement on the design of Lipinski & Nelson (1974). In addition, it is clear that the administration of polygraph examinations was associated with increased reliability. The results may be analogous to those of Lipinski & Nelson (1974) in that the reliability of self-recording in their study increased when subjects were aware of overt reliability assessment. The chief advantage of the polygraph procedure is that one need not have access to the performance of the behavior itself in order to maintain consistency in and increase the reliability of self-recording.

It appears, then, that polygraph examinations may be useful in those therapeutic situations where the therapist has access only to the subject's self-report of the behavior. This is not to say that one now has an indication of reliable data, but rather that the probability of a subject's recording reliably and consistently is increased. Further, the finding that in the polygraph sessions the reactive effect of self-recording was facilitated suggests that reliable self-recording, induced by polygraph examinations, will enhance
the reactive effects (therapeutically) desired. Finally, since the Polygraph technique does not require that the subject receive feedback on the examination (Reid & Inbau, 1966), the necessity of professional assistance is obviated. Thus, the procedure might be utilized by any therapist with access to a polygraph to increase both the reliability and the reactivity of a client's self-recording.

In spite of the possibilities of this procedure in practice, the results of this study need be interpreted with caution. First, there is no data to suggest that the effects of polygraph examinations on self-recording will generalize to the natural environment. Second, as has already been mentioned, additional sessions might have clarified a few points. Further, additional polygraph sessions might answer the question as to whether the maintenance of consistent, reliable self-recording by polygraph examinations attenuates with time. It seems that such might be the case. The ultimate solution would probably be to give polygraph examinations on a variable schedule. This type of schedule ought to maintain consistent, reliable self-recording over extended periods of time. Further research into the usage of polygraph examinations is therefore in order. Nevertheless, the results indicate that the polygraph examination is a useful tool suitable for clinical application at this time.
REFERENCES


Fox, L. Effecting the use of efficient study habits. Journal of Mathematics, 1962, 1, 76-86.


Hawkins, R. & Dotson, V. Reliability scores that delude: An Alice in Wonderland trip through the misleading characteristics of inter-observer agreement scores in interval recording. Paper presented at
the Third Annual Symposium on Behavior Analysis, Lawrence, Kansas, May 10, 1972.


APPENDIX A

Name ___________________  Age ____________
Address ___________________  Phone ____________
Year in school _________  Major ____________

Please answer the following questions:

1. Approximately how many hours per week do you watch T.V.?

2. Do you only watch certain programs? If so, which ones?

3. Where do you usually watch T.V.?

4. While watching T.V., do you often:
   - play cards?
   - eat? (if so, meals or snacks?)
   - talk with friends?
   - read?
   - study?
   - other? (what?)

5. Do you usually watch T.V. alone? If not, then with whom?

6. What psychology courses have you taken or are you now taking?

7. Have you had any experience in recording behavior?