VALIDATION OF AN INSTRUMENT TO MEASURE AN INTERVIEWER'S ABILITY TO CONDUCT A 24-HOUR FOOD RECALL

By

KRISTA KAYE HAMILTON

Bachelor Of Science

Oklahoma State University

Stillwater, Oklahoma

1999

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirement for
the degree of
MASTER OF SCIENCE
December, 2003

VALIDATION OF AN INSTRUMENT TO MEASURE AN

INTERVIEWER'S ABILITY TO CONDUCT A

24-HOUR FOOD RECALL

Thesis Approved:

Kathryn S. Kamy

Thesis Advisor

Elizabeth A. Droke

Dean of the Graduate College

Acknowledgements:

I would like to thank my major advisor, Dr. Kathryn Keim for her guidance, assistance and patience. I sincerely thank my committee, Dr. Glenna Williams, Dr. Elizabeth Droke, and Dr. Maria Spicer. Thank you to Dr. Christine Johnson for design and statistical information and assistance. Thank you to Gary Johns for his assistance and recommendations in producing the videotaped interviews. Thank you to all who assisted in any way.

I would also like to thank my dad, Sam Hamilton and my husband, Brandon Neal for their encouragement and support.

TABLE OF CONTENTS

LIST OF APPENDICES

LIST OF TABLES

CHAPTER I.

INTRODUCTION

The purpose of the present study was to validate the evaluation tool that is a part of the *The 24-Hour Food Recall In-Service Training Program. The 24-Hour Food Recall In-Service Training Program* is an educational curriculum developed by Oklahoma State University Cooperative Extension Service. The program consists of an instructional video and accompanying booklet. The purpose of the curriculum is to train Nutrition Education Assistants (NEA's) to conduct effective 24-hour food recall interviews. The training includes an evaluation tool used to determine if the NEA has achieved the skills needed to conduct an effective 24-hour recall interview. This study will help to evaluate the NEA's ability to perform effective 24-hour food recall interviews. By improving interview techniques, food intake data and program evaluation data will be more accurate and valid.

A 24-hour food recall is a method to determine what a person has recently eaten. During the recall, an interviewer asks the respondent questions to help the respondent remember everything the respondent ate in the previous 24-hours. Twenty-four hour recalls are often used to estimate nutrient intake for populations. If multiple recalls are conducted for one respondent, the recalls can also estimate nutrient intake for that individual. If using information from multiple recalls, the interview is one part of nutrition assessment.

The quality of the data obtained during a recall interview is dependent on the interviewer's skill. Training and monitoring interviewers is crucial to the validity of the data obtained in the interview (Barrett-Connor, 1991) because a skilled interviewer can minimize misreporting (Williams, 1997). A major weakness of the 24-hour recall is the misreporting of dietary intake (Barrett-Connor, 1991; Mela and Aaron, 1997). Misreporting can be minimized by building rapport between the interviewer and respondent and by asking questions to help the respondent remember what was eaten (Williams, 1997).

There are multiple reasons to validate the evaluation tool. A valid tool is necessary to reduce subjectivity in evaluating interviewers. By using a valid tool evaluators can only rate given characteristics, ignoring characteristics which may have no bearing on the validity of the interview data. When using a tool each skill can be weighed evenly. Assessing the validity of the tool will help to be certain it measures the desired characteristics. Because training and monitoring interviewers is crucial to increase the accuracy of the recall it is necessary to have an appropriate method of evaluating interviewers' skill. If the tool is not valid, it cannot be used to adequately determine interviewers' skill level.

This study had three phases. In phase one, the literature was reviewed to determine what characteristics are necessary to perform an effective 24-hour recall. Any characteristic mentioned at least three times was included in a primary trait analysis (PTA). In phase two, the PTA and desired characteristics were used to create three videotaped interviews using good, fair or poor interview techniques. To validate the videotaped interviews, subjects then used the PTA to make sure the interviews were at

the intended quality level. For phase three, experts used the Evaluation of Food Recall Techniques (EFRT) tool to evaluate the interviews. To determine intrarater reliability, each expert evaluated one interview twice.

The three hypotheses:

- Hypothesis 1: Scores on the PTA will differ significantly and be highest for the good interview, in the mid-range for the fair interview, and lowest for the poor interview.
- Hypothesis 2: Scores on the EFRT will differ significantly and be highest for the good interview, in the mid-range for the fair interview, and lowest for the poor interview
- Hypothesis 3: For each expert, the scores of the first viewing of each interview will differ from the scores on the second viewing of the same interview.

CHAPTER II.

REVIEW OF LITERATURE

This review of the literature will discuss the advantages of the 24-hour food recall, the expense involved with the 24-hour recall, disadvantages of the recall, the importance of interviewer training, and assessing the validity of an interviewer evaluation tool.

Advantages of the 24-hour Food Recall

Ease of administration is one advantage of 24-hour recall. The 24-hour recall does not require the participant to be able to read English (Thompson and Byers, 1994), because the interviewer verbally asks questions and writes the answers. However, the interviewer and respondent should be able to fluently speak the same language (IOM, 2000). The recall can be done quickly; an acceptable 24-hour recall can take approximately 20 minutes (Lee and Nieman, 1996; Thompson and Byers, 1994; Wolper et al, 1995). The recall does not require the respondent to keep lengthy food records (Williams, 1997; IOM, 2000).

Another advantage of the 24-hour recall is minimal respondent burden. Other dietary assessment methods, such as weighed food records, observed intake, or multiple-day diet records carry a large amount of perceived work for the respondent. The respondent does not need any training or preparation prior to doing a 24-hour food recall

interview. Respondents are more likely to participate in a study if the method of dietary assessment requires minimal respondent burden (Lee and Nieman, 1996; Thompson and Byers, 1994; Wolper et al, 1995).

The 24-hour recall interview relies on short-term memory (Thompson and Byers, 1994; Wolper et al, 1995). This is an advantage of the 24-hour recall because most people cannot remember what they ate more than 24 hours previously (Baranowski and Domel, 1994)

Because the 24-hour recall is retrospective, there is less chance of change in eating behavior (Lee and Nieman, 1996; Thompson and Byers, 1994; Wolper et al, 1995). Altering intake for positive self-presentation or to decrease the amount of food to weigh or record is the Guinea Pig Effect (Frank et al, 1977; Windsor et al, 1994). In the Guinea Pig Effect, the act of measurement is a change agent because knowing a measurement is going to occur may lead to a change in behavior. Not knowing when the recall will occur means there will not be a Guinea Pig Effect.

Compared to other methods of estimating dietary intake, the 24-hour recall can produce more complete data. During a 24-hour food recall, everything a respondent eats or drinks can be recorded. In the 24-hour recall, the interviewer can ask specific questions about each food or beverage to ascertain specific information, such as how the food was prepared or what might have been added during or after cooking. If the respondent is using a generated list of foods, such as in a food frequency, information about the actual foods consumed may be missing or incorrect. In a food frequency questionnaire, only the foods listed can be reported (IOM, 2000). In any food intake recall method, including food frequency questionnaires and 24-hour recall interviews,

when a respondent is asked a question either verbally or on paper, the respondent will give one of three options, no answer, a wrong answer, or a correct answer (Kohlmeier, 1994). In a face-to-face or telephone 24-hour food recall there is less chance of no response because the interviewer should pursue the question until an answer is given.

Expense of the 24-hour Recall Interview

There is disagreement as to the expense involved in performing 24-hour recall interviews. In contrast to Lee and Neiman (1996) and Frank et al (1977), some researchers believe the 24-hour recall is expensive in terms of time and money (Barrett-Connor, 1991; Buzzard, 1994; Windsor et al, 1994). The manpower and financial cost of the 24-hour recall can be considered both an advantage and disadvantage, depending on perspective. Because weighed food records require extensive training and weighing equipment for each respondent, the 24-hour recall method is considerably less expensive than weighed food records (Lee and Nieman, 1996). However, effective 24-hour recall interviews require extensive time to train the interviewer (Barrett-Connor, 1991), which can increase cost. When collecting data on a large sample, a relatively small number of trained staff can collect data (Frank et al, 1977). Training and monitoring interviewers is crucial to increase the accuracy of the recall and increases the cost (Barrett-Connor, 1991; Buzzard, 1994; Windsor et al, 1994). Time and money is wasted when an inaccurate interview is conducted because the inaccurate interview creates unusable data. The proportion of interviewers to respondents is the determining factor in deciding the expense of conducting the interviews.

Disadvantages of the 24-hour Recall Interview

The most prominent disadvantage to the 24-hour food recall is inaccuracy in reporting (Barrett-Connor, 1991; IOM, 2000; Lee and Neiman, 1996; Mela and Aaron, 1997; Thompson and Byers, 1994; Wolper et al, 1995). When compared to observation of intake, Karvetti and Knuts (1985) found there are three key sources of inaccuracies in reporting: omissions, additions and misidentification. Omissions include not reporting foods that were consumed and under reporting of portion size. Additions include reporting foods that were not consumed and over reporting of portion size. Misidentification occurs when the food the respondent consumed is not what food the interviewer records. Unknown foods, such as foods from ethnic restaurants, also lead to misidentification in reporting (Baranowski and Domel, 1994; Kohlmeier, 1994).

Because every respondent's behavior is based on different motivations, there are many intentional reasons for misreporting that can make the data obtained from the recall inaccurate (Kohlmeier, 1994). The three primary reasons the respondent may intentionally misreport intake are the question or interviewer is threatening, the respondent wants to give a "desirable" answer, or the respondent wants to give a positive self-presentation (IOM, 2000; Kohlmeier, 1994).

There are several reasons for unintentional misreporting of dietary intake. One source of unintentional misreporting is lapse in memory or inaccurate memory of what was eaten (Kohlmeier, 1994). Most adults cannot visually estimate quantity and will therefore, incorrectly estimate portion sizes of foods eaten (Baranowski and Domel,

1994; Howat et al, 1994). Another source of unintentional misreporting is decreased respondent concentration (Kohlmeier, 1994). Decreased respondent concentration can include fatigue, inattention or apathy toward the interview as well as toward the act of eating (Baranowski and Domel, 1994; Kohlmeier, 1994). Decreased interest in eating due to mood, or concurrent activities also leads to underreporting (Baranowski and Domel, 1994). Another source of unintentional misreporting is the Flat Slope Syndrome (Baranowski and Domel, 1994). The Flat Slope Syndrome states that as intake increases, either by larger portion sizes, or a large number of foods eaten, people tend to underestimate intake of foods (Baranowski and Domel, 1994; Karvetti and Knuts, 1985). As intake decreases, over reporting increases (Baranowski and Domel, 1994; Karvetti and Knuts, 1985).

One 24-hour recall interview cannot show usual intake. Similarly, one 24-hour recall interview on one day cannot show seasonal variations of individual intake (IOM, 2000; Lee and Nieman, 1996; Thompson and Byers, 1994). However, an interview on one occasion can describe the usual portion sizes of the foods eaten by an individual on the day of measurement (Kohlmeier, 1994).

The 24-hour recall interview is not appropriate for all participants. Because of the need to rely on short-term memory, the 24-hour recall is not appropriate for anyone with cognitive deficits (Baranowski and Domel, 1994; Van Staveren et al, 1994; Wolper et al, 1995). In addition, it is a disadvantage if the interviewer and the respondent do not fluently speak the same language (IOM, 2000; Wolper et al, 1995). Because respondents tend to build better rapport with interviewers who they perceive as similar to themselves, differences between interviewer and respondent, such as gender and perceived

socioeconomic status, may cause variations in responses (IOM, 2000; Windsor et al, 1994 Wolper et al, 1995).

Because interviewer training is important to lessen the disadvantages of the interview (Wolper et al, 1995) ineffective interviewer training can be viewed as a disadvantage. *The24-Hour Food Recall In-Service Training Program* is designed to train former program participants to perform interviews, to reduce differences in demographics of the interviewer and respondent.

Interviewer Training to Decrease Disadvantages

The value of the 24-hour recall interview and the data it produces is dependent on the interviewers' skills (Barrett-Connor, 1991). To perform effective 24-hour recall interviews interviewers should be extensively trained (Barrett-Connor, 1991; McNutt et al, 1998) and should have frequent feedback from trainers after the data collection process has begun (McNutt et al, 1998). If interviewers are not adequately trained, the data from one respondent on one date is difficult to reproduce when interviewed by more than one interviewer (Frank et al, 1977; Wolper et al, 1995). To reduce measurement error between interviewers, interviewers should be trained by a standardized protocol (Frank et al, 1977; Lee and Nieman, 1996).

A skilled interviewer can minimize misreporting by building rapport with the respondent (McNutt et al, 1998). Clearly explaining the purpose of the interview, speaking the same language as the respondent (IOM, 2000), and being patient with the

respondent are ways the interviewer can build rapport with the respondent and decrease error.

Interviewers must be trained to ask appropriate open-ended and probing questions to help the respondent remember additions (such as salt and butter), beverages and snacks (Kohlmeier, 1994; Thompson and Byers, 1994). An example of a probing question is "what did you add to your food at the table?" Without probing respondents tend to omit additions and beverages (Lee and Neiman, 1996). If the respondent is elderly, fading memory requires the interviewer to be especially diligent in asking probing questions (Van Staveren et al, 1994). Interviewer probing can increase reported intake by 25% (Thompson and Byers, 1994).

Responses carry perceived social desirability (Windsor et al, 1994; Wolper et al, 1995), and respondents generally like to give a positive self-presentation (Baranowski and Domel, 1994; Kohlmeier, 1994). A skilled interviewer can reduce this by not showing approval or disapproval of responses during the interview (Johnson et al, 1998; Lee and Neiman, 1996; Wolper et al, 1995).

Unintentional misreporting is somewhat predictable and can be easily addressed by a skilled interviewer. An interviewer who is trained to remain focused can reduce error caused by decreased respondent concentration (IOM, 2000; Williams, 1997).

Interviewers must be trained to help respondents estimate the amount of foods eaten. Because most adults find it difficult to visually estimate quantity (Baranowski and Domel, 1994; Howat et al, 1994; IOM, 2000), using two and three dimensional models, household measures and utensils can help the respondent more accurately estimate portion size (Howat, 1994; Jonnalagadda et al, 2000; Karvetti and Knuts, 1985; Lee and

Neiman, 1996; McNutt et al, 1998; Wolper et al, 1995). When a respondent can see a visual representation of a specific quantity of food, they can more accurately compare their intake with the provided representation.

The 24-Hour Recall In-Service Training Program is a program created by Oklahoma State University Extension to train Nutrition Education Assistants (NEA'S), who are not nutrition professionals, to conduct effect 24-hour food recall interviews (Williams, 1997). This program uses standardized training materials to assist the interviewer in conducting an effective 24-hour recall.

Assessing the Validity of an Evaluation Tool

Assessing the validity of a 24-hour food recall interviewer evaluation tool is necessary to be certain the evaluation tool has measured the desired characteristics based on the objectives of the original education program. The tool must be properly designed to correctly assess the desired characteristics (Windsor et al, 1994). If the evaluation tool is not valid, it will not accurately measure that which it was designed to measure.

The present study was designed to assess the content validity and criterion validity of *The 24-Hour Food Recall In-Service Training Program* Evaluation of Food Recall Techniques (EFRT) tool. For the tool to have content validity it should have items that measure all characteristics that are considered important to perform a thorough 24-hour food recall interview (Schutt, 1999; Vogt, 1999; Windsor et al, 1994). Content validity is based on expert opinion and the research literature. Content validity is assessed by identifying interviewer characteristics necessary to conduct an effective

recall interview and developing an instrument to measure the characteristics. For the tool to have criterion validity it should measure similar results when compared to an acceptable criterion (Schutt, 1999; Vogt, 1999; Windsor et al, 1994). In the present study, the criterion against which the EFRT was compared was the three videotaped interviews created representing three skill levels of recalls made after conducting a primary trait analysis (PTA). The PTA can be considered more accurate than the EFRT because the PTA assesses desirable characteristics that were identified in the literature. This study used concurrent criterion validity. The PTA and EFRT were not used at the same time, but were used to measure the same videotaped interviews.

CHAPTER III.

METHODOLOGY

The purpose of the present study was to assess the reliability and validity of the evaluation of food recall techniques (EFRT) tool that is part of *The 24-Hour Food Recall In-Service Training Program* developed by Oklahoma State University Cooperative Extension. The EFRT tool is used in evaluating the ability of a paraprofessional to perform an effective 24-hour food recall interview. For the EFRT tool to have content validity it should address all traits that are considered necessary to perform an effective 24-hour recall interview (Schutt, 1999; Vogt, 1999). The traits needed to perform an effective 24-hour food recall interview were determined by examining previous research. All aspects of the study were examined and approved by the Institutional Review Board of Oklahoma State University. The IRB approval form and continuation approval forms are in Appendix A and B.

Research Design

This study had three phases. In phase one, the literature was reviewed to determine what characteristics are necessary to perform an effective 24-hour recall. Any characteristic mentioned at least three times was included in a primary trait analysis (PTA). In the second phase, the PTA and desired characteristics were used to create three videotaped interviews using good, fair or poor interview techniques. To validate the videotaped interviews, subjects then viewed the videotaped interviews and evaluated

the quality of the interview using the PTA. The PTA was used to ensure the videotaped interviews were at the intended quality level. For the third phase, a second population of subjects or experts was selected to validate the evaluation tool. The experts viewed the videotaped interviews and used the Evaluation of Food Recall Techniques (EFRT) tool to evaluate the interviews. To determine intrarater reliability, each expert evaluated one interview twice. The non-experimental, correlational research design used a self-administered evaluation tool and videotapes of 24-hour food recall interviews to test the hypotheses.

The three hypotheses:

- Hypothesis 1: Scores on the PTA will differ significantly and be highest for the good interview, in the mid-range for the fair interview, and lowest for the poor interview.
- Hypothesis 2: Scores on the EFRT will differ significantly and be highest for the good interview, in the mid-range for the fair interview, and lowest for the poor interview.
- Hypothesis 3: For each expert, the scores of the first viewing of each interview will differ from the scores on the second viewing of the same interview.

Primary Trait Analysis

There were four purposes of the primary trait analysis (PTA). One purpose was to define explicit criteria to be used in creating three videotaped interviews. Reviewing the research literature identified the criteria. Any characteristic mentioned at least three

times was included in the PTA. Characteristics on the PTA were used when filming three videotaped interviews of varying interviewer skill level. Another purpose of the PTA was to reduce subjectivity in evaluating the interviewer's skill (Walvoord and Anderson, 1998). By using the PTA, the evaluators could only rate the interviewer's skill on given characteristics. Without the PTA the evaluator would be able to rate the interviewer based on any characteristic, such as physical appearance, which may have no bearing on the validity of the interview data. In addition, by using the PTA each skill was weighed evenly. A third purpose of the PTA was to make sure the videotaped interviews were at the intended quality level. Subjects viewed the interviews and used the PTA to evaluate the interviewer's skill. The fourth purpose of the PTA was to create a numeric scale for evaluating each videotaped interview.

In creating the PTA for this study a list of characteristics were needed. The characteristics were skills needed to conduct an effective 24-hour recall. Reviewing the research literature identified the characteristics. As the literature was read, a matrix of desirable characteristics was produced. Each time a study mentioned a characteristic necessary for an effective 24-hour recall, the characteristic it was tallied on the matrix. Any characteristic mentioned in at least three sources was used for the PTA.

Eighteen characteristics were identified and used for the PTA. The interview should cover 24 Hours. The interviewer should introduce self, explain the purpose of interview, and explain the importance of reporting actual intake. Interviewer should not verbally or nonverbally indicate any approval or disapproval of the respondent's answer. Interviewer should guide the respondent through the interview using open-ended questions, without suggesting responses. Interviewer should use verbal questions and

visual cues and ask about activities performed during the day to help respondent remember what foods were eaten. Interviewer should obtain estimates of the portion size of foods that were eaten. Interviewer should use two-dimensional visual aids and three-dimensional visual aids to help the respondent visualize the type of food and portion sizes consumed. The models should realistically represent food or portion sizes. Interviewer should use measuring cups, measuring spoons, common size plates, bowls or drinking glasses to help the respondent estimate portion size. Interviewer should ask the specific name of foods that were eaten, and clarify that the interviewer and respondent are visualizing and conceptualizing the same food. Interviewer should obtain brand names of foods that were not prepared from a home recipe, including names of restaurants where foods might have been purchased. Interviewer should ask how foods were prepared and ask what cooking methods (bake, fry, raw) were used including asking what was added to foods before eating. Interviewer should keep interview focused and on track and avoid meal labeling unless done by the respondent.

For the PTA, most characteristics were given a four-point scale. The points were "Interviewer usually performs", "Interviewer occasionally performs", "Interviewer rarely performs", "Interviewer does not perform". Four characteristics were evaluated using yes or no questions. To score four points, the interviewer had to usually perform the skill or behavior to conduct an effective 24-hour recall. If the interviewer occasionally performed a skill the interviewer scored three points. If the interviewer rarely performed a skill, the interviewer scored two points. If the interviewer did not perform a skill, the interviewer received one point. For the 'Yes" or "No" questions four points were given for a yes answer, and one point for a no answer.

There were 18 characteristics on the PTA. Because the format of the first question on the PTA ("Covers 24 hours") was confusing, many subjects did not answer it. For most of the characteristics the skill level words (for example, "Interviewer usually performs") were above the rows of questions and subjects were to indicate their choice by making a mark in a box. For "Covers 24-hours" the words "yes" and "no" were on the same row as the question and subjects were to circle their choice. This format imitated the format of the Evaluation of Food Recall Techniques tool. The problem was not corrected, because the all subjects had returned all materials before the problem was noticed. The question "Covers 24 hours" was not included in any statistical analyses.

After throwing out one question, the minimum score was 17 and the maximum score was 68 and created by summing the value for each item. A score of 68-51 points on the PTA was considered a good interview, 50-34 points for a fair interview, and 33-17 points for a poor interview. The PTA form is in Appendix C.

Development of Videotaped Interviews

After the PTA was completed three scripts and videotapes were developed. The purpose of the videotaped interviews was to have a consistent and uniform interview for all subjects to view. The subjects in the second phase would validate that the interviews were at the intended quality level. The subjects in the third phase would use the interview of predetermined quality to validate the Evaluation of Food Recall Techniques tool. Each script and videotaped interview depicted one of three interviews: a good (PTA score 68-51), a fair (PTA score 50-34), and a poor (PTA score 33-17). The actors were a professional actress (the respondent) and the primary researcher (the interviewer). The actress was employed through a professional talent agency. Because the actress was

paid, she did not sign a consent and release form. The interviews were videotaped in a private home and the actress and researcher wore their own clothing and did their own hair and makeup. The actress was told her recall should include one meal at a restaurant and microwave popcorn; she could chose any other foods she wished. The script the actress was given is Appendix D. She was encouraged to report the foods she actually ate. She was not to reveal any information that was not solicited by the interviewer. The poor interview was filmed first, followed by the fair interview, filming the good interview last. They were filmed in this order because as the interviewing skills increased the actress would be aware of more detailed information about her foods. Because the lower quality interviews were filmed first, she only knew the foods she ate and approximately what portion sizes; she could not accidentally volunteer information that was not uncovered by the interviewer. The videotaped interviews included time when the interviewer was writing responses. This time increases with each interview. It was important to show the interviewer writing responses, because that can be considered part of the interview.

The good interview was approximately 17 minutes in length. The fair interview was about 7 minutes and poor interview was about 3 minutes in length. The videotape labels and video jackets were color-coded and forms were color coded to correspond with the videotapes. This was done to lessen confusion and to make it easy for the researchers to see which form should be sent with each tape. The good interview was coded "Q" and colored orange. The fair interview was coded "F", colored blue. The poor interview was "W", coded black. The letter and number codes were selected to reduce the possibility of

bias due to common letters such as A, B, C or the colors red, yellow, green for traffic light color schemes.

Validation of Videotaped Interviews

To validate whether the created videotaped interviews accurately reflected the three quality levels of 24-hour food recall interviews, undergraduate and graduate students from Oklahoma State University Department of Nutritional Sciences validated the videos.

Subjects

The subjects were recruited through two courses, NSCI 4853, Medical Nutrition Therapy I and NSCI 5713, Community Dietetics. These two courses were selected because their curricula included training in performing and evaluating 24-hour food recall interviews. One male and 50 females were recruited as subjects. The response rate was 100% of students in each course. Receiving extra credit points compensated subjects. Students could choose not to participate and complete an alternative project for the same number of extra credit points. All subjects signed a consent form (Appendix E).

Methods

Fifty-one subjects chose to participate. The subjects were given written instructions. The written instructions gave the order in which the interviews were to be viewed. There were three different instruction sheets, each giving a different order in which to view the

interviews (Appendices F through H). The subjects were randomly assigned the order to view the interviews. There was no method to make sure the videotaped interviews were viewed in the given order. The subjects watched each interview and scored the actions of the interviewer with the PTA criteria sheet (Appendix C).

Statistical Analysis

- Hypothesis 1: Scores on the PTA will differ significantly and be highest for the good interview, in the mid-range for the fair interview, and lowest for the poor interview.

A score was created for each subjects' PTA tool by summing the value for 17 items on the PTA tool, and a mean generated for each interview. Cronbach's alpha was used to determine the internal reliability of the scale across the three interviews. The general linear model method was used to determine if the mean score on each videotaped interview was significantly different among all three interviews. Spearman's rank correlation was used to determine the correlation between the PTA mean scores on the three videotapes. This was to determine if scales for each interview were related. Results were analyzed using SPSS version 10.0 with significance at $p<0.05$.

Validation of Evaluation Tool

Subjects and Recruitment

The subjects or experts were members of the Food and Nutrition Extension Educators and Higher Education Divisions of the Society for Nutrition Education. The

names and addresses of members of the two divisions were purchased from the Society for Nutrition Education for $100. After removing members with foreign addresses and duplicates between the two lists, there were 245 members of the Food and Nutrition Extension Educators Division and 137 members of Higher Education Division. Two members were removed because they were faculty of the Nutritional Sciences Department at Oklahoma State University, leaving 135 Division of Higher Education members. The total population was 380.

The sample was defined as a purposive sample because potential subjects were selectively recruited. The first group of subjects was recruited by sending a letter to a random selection of 200 members of both divisions. Due to the inadequate number of subjects from the first 200 selected, recruitment packets were sent to the remaining 180 members of each division using the same procedures as the original recruitment packet. The packet contained a letter that asked them to participate in the study, described the purpose of the study and how recipients were selected, and addressed confidentiality and voluntary participation (Appendix I). Included with the letter was a schedule of participation (Appendix J). Demographic information (Appendix K) was collected to be able to describe the subjects. Subjects signed an informed consent form (Appendix L). All volunteers who met the inclusion criteria were accepted for the study. The inclusion criteria included membership in either Division, returning a signed consent form and completing all study materials. However, three subjects were not members of Society for Nutrition Education, but mid-level managers for the Expanded Food and Nutrition Education Program on the state or county level. They were recruited after the person to whom the mailing was sent passed on the recruitment packet. Estimated minimum time

involvement for the subjects was 12 weeks with two to five total hours to view the interviews. At the end of the study subjects were compensated by receiving the three videotapes.

Distribution of Videotaped Interviews and Evaluation Tools

Each expert viewed three interviews in random order, and viewed one interview twice to determine intra-observer reliability. The order of viewing was determined for each subject by using a repeated cycle. The repeated video each subject viewed was determined by including the repeated interview into the order cycle (Appendix M). Each subject number had an assigned viewing order, and each subject was assigned a subject number as they returned enrollment materials. Experts were instructed to watch each videotaped interview. After viewing each interview, the experts rated the interviewer's food recall skills using the Evaluation of Food Recall Techniques (EFRT) tool, developed by Sheik and Williams (Appendix N). With each videotaped interview we sent an instruction letter (Appendix O) and a color-coded copy of the EFRT tool. The color of the paper used for the EFRT tool matched the color of the video jacket and the same colored dot on the videotape. Each expert viewed the interviews in one of several orders. The subjects only had one videotaped interview in their possession at any time. As completed materials were returned, the next videotape and EFRT tool were mailed. If subjects did not return materials within 15 mailing days, a reminder postcard was sent (Appendix P). A second reminder postcard was sent 15 mailing days after the first reminder postcard.

Evaluation Tool Scores

The EFRT tool is presented in Appendix N. Glenna Williams and Brenda Sheik created the form. The EFRT tool consisted of twelve questions concerning the quality of the interviewers skills. However, for this study only nine questions were chosen for analysis. The question "Did food recall cover a 24-hour period?" was not included in any statistical analyses because the format of the form was confusing and many subjects did not answer the question. For most of the characteristics the skill level words (for example, "Needs improvement") were above the rows of questions and subjects were to indicate their choice by making a mark in a box. For "Covers 24-hours" the words "yes" and "no" were on the same row as the question and subjects were to circle their choice. The problem could only be corrected by altering the format of the tool. If the format were altered it would not be the same as the tool being validated. Because so few subjects answered the question, "Covers 24 hours" was not included in any statistical analyses. One question asked if forms were completed properly. The forms were not considered part of this study, so the question "Was Food Recall form properly and thoroughly completed?" was excluded. Another excluded question was "Was overall Food Recall accurate?" Subjects did not have the actual intake data available to them, and therefore were unable to assess the accuracy of the interview. These three questions were ignored in all analyses.

After removing the three questions, scores on nine items were analyzed. When the experts evaluated each interview, summing a value from each characteristic on the

evaluation tool created a composite score. The rating on each item received a score ranging from one to four. One meant "Needs Improvement", two meant "Fair", three meant "Good", and four meant "Good Job!". The highest possible score was 36, while the lowest possible score was 9. Summing the values for each item generated the score. After the first wave of EFRT tools was mailed out, a mistake was noticed on the EFRT tool. For the first viewing one question was omitted from the EFRT tool. When the EFRT tool was used for the second, third and fourth viewings the question was included. For statistical analysis, a value was generated for the omitted question by averaging the score on that question for each of the three individual videotapes.

There were three levels of interview quality on three different videotapes: good, fair, and poor. Scores on the evaluation tool were a composite measurement based on ratings on nine items of the evaluation tool. The expected range of scores for each interview was 36-28 for the good interview, 27-18 for the fair interview, and 17-9 for the poor interview.

Statistical Analysis

- Hypothesis 2: Scores on the EFRT will differ significantly and be highest for the good interview, in the mid-range for the fair interview, and lowest for the poor interview.
- Hypothesis 3: For each expert, the scores of the first viewing of each interview will differ from the scores on the second viewing of the same interview.

A score was generated for each expert's EFRT by summing the values on nine EFRT items and generating a mean for each interview. Cronbach's alpha was used to determine the internal reliability of the scale across the three interviews. The general

linear model method was used to determine if the mean score on each videotaped interview differed significantly among all three interviews. Spearman's rank correlation was used to determine the correlation between the EFRT mean scores on the three videotaped interviews. This was to determine if scales for each interview were related. A paired t-test between the scores of the first and second viewings of the same videotape was used to determine if the scores differed significantly between the first and second viewing. Results were analyzed using SPSS version 10.0.

CHAPTER IV.

RESULTS

Primary Trait Analysis

The purpose of validating the videotaped interviews was to be sure the interviews were at the intended quality level. The subjects for the validation of videotaped interviews were nutritional sciences students from Oklahoma State University. The subjects were enrolled in either a senior level Medical Nutrition Therapy course, or a graduate level Community Dietetics course. These two courses were selected because the coursework included training in performing and evaluating 24-hour food recall interviews. One male and 50 females were subjects (n=51). Some subjects did not answer every question on all three primary trait analysis (PTA) forms. When analyzing data, if a PTA form was not complete, the data was not included in analysis. Therefore, the sample size is not always 51.

On the primary trait analysis (PTA) the format of the question "Covers 24 hours" was confusing, and many subjects did not answer it. For this reason, the question "Covers 24 Hours" was excluded from statistical analyses. After removing this question, there were 17 items included in the statistical analysis of the PTA tool

Cronbach's alpha score greater than 0.70 implies the scores are reliable (Vogt, 1999). Cronbach's alpha for all items from all three interviews (n=41) was 0.84. Cronbach's alpha for the poor interview (n=47) was 0.84 and for the fair interview

(n=45) was 0.73, suggesting the PTA scores for the poor and fair interviews were reliable (Vogt, 1999). For the good interview (n=48) alpha was 0.64. The reliability scores indicate the subjects were more consistent when evaluating the fair and poor interviews than when evaluating the good interview.

The expected score range for each videotape was 51-68 for the good interview, 34-50 for the fair interview, and 17-33 for the poor interview. All three mean PTA scores were within the desired ranges (Table 1). Using general linear model and ANOVA the poor, fair and good interview scores were significantly different ($p<0.05$) (Table 1).

Spearman's rho correlation (Table 2) between the poor and fair interviews was significant ($p<0.01$). This indicates the scores for the poor interview and the scores for the fair interview are not discrete. There was no correlation between the good and fair interviews or between the good and poor interviews.

The scores for each individual item on the PTA are presented in Table 3. The overall trend was lowest ratings on the poor interview, intermediate ratings on the fair interview, and highest ratings on the good interview.

Evaluation of Food Recall Techniques Tool Validation

Of the 380 members of Society for Nutrition Education Higher Education and Food and Nutrition Extension Educators Divisions, 81 subjects initially volunteered. The response rate was 21%. Only 66 subjects returned usable data (response rate 88%). Two subjects withdrew from the study, while 12 subjects did not complete the evaluation for all four videotapes. One subject signed a consent form, but never completed any

evaluations. When analyzing data, if an EFRT tool was not complete, the data was not included in analysis. Therefore the sample size is not always 66.

Demographic information of the experts is presented in Table 4. Because some questions on the demographic form allowed subjects to select more than one answer, some totals are more than 100%. The majority of experts were employed as full-time faculty with either cooperative extension or a university. The job function for most experts was nutrition education. Many experts reported having earned a Master's or Doctoral degree. The focus of continuing education for most experts was community nutrition issues (82%) or adult education issues (44%). The question "How often do you conduct 24-hour recall interviews?" was open-ended and experts could fill in any response. The responses were compiled into two categories: more often than once a week, and less often than once a week. Most experts perform 24-hour recalls interviews less than once a week. Because 74% of the experts were responsible for training others to conduct 24-hour recall interviews, the population was appropriate for evaluating the tool, even though the experts perform recall interviews less than once a week.

Cronbach's alpha for the EFRT tool for all three interview quality levels was 0.83 (n=50). For the good interview (n=60), Cronbach's alpha was 0.80. Alpha scores greater than 0.70 are considered reliable (Vogt, 1999). For the poor (n=58) and fair (n=60) interviews Cronbach's alphas were 0.54 and 0.67, respectively. The reliability scores indicate the experts were more consistent when evaluating the good interview than when evaluating the poor and fair interviews.

The second hypothesis was that when the experts evaluated the interviews using the Evaluation of Food Recall Techniques (EFRT) tool scores on the evaluation tool

would differ significantly among interviews. The scores would also be highest for the good interview, in the mid-range for the fair interview, and lowest for the poor interview. The expected ranges of EFRT scores were 28-36 for the good interview, 18-27 for the fair, and 9-17 for the poor interview. The scores were all within the expected range (Table 5). Using general linear model and ANOVA each mean score was significantly different ($p=0.000$). Individual item scores for each interview are presented in Table 6.

Using Spearman's rho correlation, the scores on the fair interview were related to the scores on the poor and good interviews (Table 7). This indicates the fair and poor interview scores were not dissimilar when compared to each other, and the fair and good interview scores were not dissimilar when compared to each other. The scores for the good and poor interviews were not correlated.

Each interviewer evaluated one interview twice to determine intrarater reliability. The third hypothesis was for each evaluator, the scores of the first viewing of each interview would differ from the scores on the second viewing of the same videotape. Using paired t-test, comparing the mean scores for the first and second viewings of the poor and fair interview did not differ significantly, Table 8. The mean scores of the good interview at the first viewing were significantly greater than the second viewing ($p<0.01$).

Rejection of Hypotheses

Hypothesis 1: Scores on the PTA will differ significantly and be highest for the good interview, in the mid-range for the fair interview, and lowest for the poor interview. We failed to reject hypothesis one because the scores on the primary trait analysis were

significantly different across all three interviews. Scores were lowest for the poor interview, intermediate for the fair interview, and highest for the good interview.

Hypothesis 2: When subjects used the EFRT tool to evaluate the interviewer's skill, scores will differ significantly, and scores on the EFRT will be highest for the highest for the good interview, in the mid-range for the fair interview, and lowest for the poor interview. We failed to reject hypothesis two because scores on the evaluation tool were significantly different. The scores were highest for the good interview, in the mid-range for the fair interview, and lowest for the poor interview.

Hypothesis 3: For each evaluator or expert, the scores of the first viewing of each videotaped interview will differ from the scores on the second viewing of the same interview. We failed to reject hypothesis three for the good videotaped interview because scores of the first viewing of the good interview differed significantly from the scores on the second viewing. We rejected hypothesis three for the poor and fair videotapes because the scores of the first viewing of the poor and fair videotaped interviews did not differ significantly from the scores on the second viewing of the same videotape.

Table 1. Primary Trait Analysis (PTA) scores of three 24-hour food recall interviews[1], n=41

	Videotapes		
	Poor	Fair	Good
PTA Score[2]	31.5 ± 7.2[a]	48.6±7.3[b]	64.7±3.5[c]

[1]Values are mean ± standard deviation.

[2] Minimum score=17, maximum score 68. The higher the score the better the quality of interview.

[a,b,c]Means with different superscripts are significantly different at $p<0.05$.

Table 2. Spearman's correlations among PTA scores of three interviews.

Videotapes	Fair	Good
Poor	0.44**	0.19
	n=43	n=45
Fair	1.00	0.22
	n=45	n=43

**Spearman's rho is significant at $p<0.01$.

Table 3. Individual PTA item scores for three 24-hour recall interviews.

PTA Item or Skill	Poor[1]	Fair[1]	Good[1]
Introduces self[a]	1.2±0.8	3.9±0.7	4.0±0.0
Explains purpose of interview[a]	1.4±1.1	2.0±1.4	4.0±0.0
Explained importance of reporting actual intake[a]	1.2±0.7	2.9±1.5	3.4±1.3
Does not show approval/disapproval[b]	2.2±0.9	2.4±1.0	3.6±0.9
Uses open ended questions[b]	2.0±0.9	2.6±0.8	3.7±0.5
Helps respondent remember what was eaten[b]	2.1±0.8	3.2±0.7	3.9±0.4
Asks about activities during the day[b]	1.5±0.8	3.0±0.9	3.9±0.4
Obtains estimates of portion size[b]	2.3±0.8	3.4±0.7	4.0±0.2
Uses 2D visual aids[b]	1.4±0.6	2.0±1.2	3.2±1.0
Uses 3D visual aids[b]	1.8±0.8	3.4±0.8	3.8±0.6
Uses measuring cups, etc[b]	1.3±0.5	2.7±1.0	4.0±0.0
Ask specific names of foods[b]	2.2±0.9	3.0±0.9	4.0±0.3
Asks specific brand names of foods[b]	2.1±1.1	2.3±0.9	3.9±0.3
Asks how foods were prepared[b]	2.4±0.9	3.0±0.8	3.8±0.6
Asks what was added before eating[b]	1.8±0.9	2.8±1.0	3.9±0.3
Keeps focused and on track[b]	3.5±0.7	3.6±0.8	4.0±0.0
Avoids meal labeling[b]	1.3±0.5	2.8±0.9	3.7±0.7

[1] Values are mean ± standard deviation

[a] 1=no and 4=yes

[b] 1=Does not perform, 2=Rarely performs, 3=Occasionally performs, and 4=Usually performs

Table 4. Demographic characteristics of experts, n=66

Demographic Characteristic		n	percent
Employment status			
	full time	55	83.3
	part time	4	6.1
	retired	4	6.1
	other	2	3.0
	not employed	1	1.5
Employment setting*			
	cooperative extension	45	68.2
	college/university faculty	30	45.5
	community/public health	6	9.1
	acute care	2	3.0
	consultation primarily to other organizations, industries or media	2	3.0
	long term care	1	1.5
	school food service	1	1.5
	commercial food service	1	1.5
	HMO, physician or other healthcare provider	1	1.5
Job functions*			
	nutrition education	56	84.8
	nutrition information/communication	32	48.5
	public health	27	40.9
	higher education	21	31.8
	research	17	25.8
	wellness/disease prevention	17	25.8
	personnel training	11	16.7
	clinical services	2	3.0
	foodservices	1	1.5
Highest degree earned			
	master's degree	32	48.5
	doctoral degree	24	36.4
	bachelor's degree	7	10.6
	associate's degree	3	4.5

Table 4. Continued

Focus of continuing education*			
	community nutrition	54	81.8
	adult education	29	43.9
	education	23	34.8
	research	14	21.2
	foods/food sciences	11	16.7
	clinical nutrition	10	15.2
	management	6	9.1
	foodservices	2	3.0
Job position			
	cooperative extension faculty at the state level	26	39.4
	cooperative extension faculty at the county level	18	27.3
	resident university faculty	16	24.2
Are you responsible for training personnel to do 24-hour food recall interviews?			
	yes	49	74.2
	no	17	25.8
How often do you conduct 24-hour food recall interviews?			
	less than once a week	48	72.7
	more than once a week	18	27.3

*Total is more than 100% because question on the demographic form allowed subjects to select more than one answer.

Table 5. Evaluation of Food Recall Techniques (EFRT) tool scores after viewing three 24-hour food recall interviews[1], n=50

	Videotapes		
	Poor	Fair	Good
EFRT Score [2]	13.6 ± 2.5[a]	20.2 ± 3.8[b]	31.9 ± 3.6[c]

[1] Values are mean ± standard deviation

[2] Minimum score=9, maximum score=36. The higher the score, the better the characteristic was performed.

[a,b,c] Means with different superscripts are significantly different at p=0.000.

Table 6. EFRT tool individual item score after viewing three 24-hour food recall interviews.

Characteristic	Videotapes		
	Poor n=58	Fair n=60	Good n=60
Did NEA prepare homemaker?	1.2±0.4	1.9±0.8	3.2±0.9
How well did NEA avoid using approval or disapproval?	1.2±0.5	1.0±0.2	3.4±0.6
Did NEA avoid getting "off track"	3.2±0.7	3.2±0.8	3.8±0.4
Did NEA use proper tools and visuals?	1.3±0.5	2.9±0.9	3.5±0.8
Did NEA ask open-ended questions?	1.7±0.7	2.3±0.8	3.5±0.7
Did NEA relate food to activity?	1.3±0.6	2.5±0.9	3.7±0.6
Did NEA include added ingredients?	1.3±0.5	2.8±0.8	3.7±0.5
Did NEA avoid meal labeling?	1.1±0.4	1.8±0.9	3.5±0.7
Did NEA avoid leading questions?	1.4±0.6	1.6±0.7	3.4±0.7

1=Needs improvement, 2=Fair, 3=Good, 4=Great job!

Table 7. Spearman's correlations among EFRT tool scores after viewing three 24-hour food recall interviews.

Videotape	Fair	Good
Poor	.52**	.22
	n=52	n=54
Fair	1.0	.46**
	n=60	n=56

** Spearman's rho is significant at $p<0.1$.

Table 8. Intrarater reliability of experts based on viewing same videotaped 24-hour food recall interview twice[1].

Viewing	Videotape		
	Poor n=16	Fair n=21	Good n=10
First	13.2±1.9[a]	19.5±4.2[a]	32.0±3.2[a]
Second	13.6±2.2[a]	20.4±3.8[a]	28.5±3.4[b]

[1] Values are means ± standard deviation

[a,b] Means with different superscripts by viewing time are significantly different using paired t-test at $p<0.01$

CHAPTER V

DISCUSSION

The purpose of this study was to determine if the Evaluation of Food Recall Techniques (EFRT) tool could be used to accurately and reliably measure an interviewer's ability to conduct an effective 24-hour food recall. The validity of the EFRT was important because training and monitoring interviewers is crucial to increase the accuracy of the recall (Barrett-Connor, 1991; Buzzard, 1994; Windsor et al, 1994). If the EFRT is not valid, it cannot be used to adequately determine interviewers' skill level in conducting a 24-hour food recall. For this study we used two tools, the EFRT created by Glenna Williams and Brenda Sheik, and the primary trait analysis tool produced for this study.

The purpose of conducting the primary trait analysis was to have a comparison to determine what characteristics were needed for the EFRT tool to have content validity. Content validity was determined by identifying characteristics needed to conduct an effective 24-hour food recall interview and developing a primary trait analysis (PTA) tool to measure those characteristics. The PTA was also necessary for determining criterion validity of the EFRT tool. To determine criterion validity scores of the EFRT tool would score the interviews in a similar manner as the PTA. In addition, the two tools needed to be used for evaluating identical interviews of known interviewer skill quality, and the PTA was used to make sure the videotapes did portray the desired level of interview skills.

The videotaped interviews were necessary for two reasons. When using two tools to determine criterion validity, the EFRT and PTA needed to be used to evaluate identical interviews. In addition, when validating the EFRT tool, all of the experts needed to evaluate identical interviews. By creating the videotaped interviews, we controlled the variable of interview quality and produced uniform, consistent interviews.

While filming the fair interview the interviewer made an obviously judgmental statement. The respondent reported eating a chicken sandwich from McDonald's. The interviewer asked if the respondent had a grilled or fried chicken sandwich. When the respondent answered, "fried" the interviewer replied, "Yuck, didn't you hear those stories about the chicken head getting fried in the chicken McNuggets? Don't eat the fried chicken at McDonald's." After this statement the respondent was much more distant than earlier in the interview. The interviewer's statement could not be edited out of the videotape, because without it, a person viewing it would not understand why the respondent was suddenly so distant. This disapproving statement may have affected the results of the study.

For the PTA, the reliability for the poor and fair interviews was greater than the reliability for the good interview. This implies that when using the PTA the subjects were more consistent when evaluating the interviewer's skills in the poor and fair interviews than when evaluating the interviewer's skills in the good interview. This could have been caused by fatigue while watching the good videotaped interview (17 minutes), as the good interview was ten minutes longer than the fair interview and 14 minutes longer than the poor interview. Each videotaped interview included portions when the interviewer was writing down responses. As the interview quality increased,

the length of time spent writing increased, because the interviewer acquires more information to record. This writing time increased the length of each interview and was tedious to watch. However, it was important to show the interviewer writing the responses since it is considered part of the interview. This was an important finding since in a real-life situation a good 24-hour recall interview should take 20 minutes or more to conduct (Lee and Nieman, 1996; Thompson and Byers, 1994; Wolper et al, 1995).

PTA scores for the poor interview and scores for the fair interview correlated (Table 2). The judgmental statement made by the interviewer during the fair interview be the reason. This statement may have affected how that particular interview was evaluated by experts and made it difficult to differentiate between the fair and poor videotapes. This may have lowered the reliability and validity of the fair videotape.

Even though the reliability for the good interview was relatively low and the fair and poor scores were correlated, the scores did differ significantly and scores were in the expected ranges. Scores were lowest for the poor interview, intermediate for the fair interview, and highest for the good interview. The scores indicated that the videotaped interviews were at the intended quality level and could be used for the EFRT validation study.

For the EFRT tool the reliability for the good interview was greater than the reliability for the fair and poor interviews. The reliability scores indicated the experts were more consistent when evaluating the good interview than when evaluating the poor and fair interviews. These findings are in contrast with the reliability for the PTA tool.

When using the PTA the reliability was lowest for the good interview, but when using the EFRT reliability was highest for the good interview. For all three interviews

the reliability of the primary trait analysis (PTA) was higher than the reliability of the EFRT. The PTA was used to measure an interviewer's skill on 17 characteristics, while the EFRT was used to measure only nine characteristics. As the number of characteristics measured increased, the reliability increased.

When using the Evaluation of Food Recall Techniques (EFRT) tool, scores from the fair interview were correlated with scores from the poor interview and scores from the fair interview were correlated with scores from the good interview (Table 7). These results suggest that when using the EFRT it was difficult for the experts to differentiate between the fair interview and the other interviews. However, the function of the fair interview was to be at a quality level between the poor and the good interviews. Fortunately, the poor and good EFRT scores were not correlated, indicating that the experts did not have a problem differentiating between the two.

The second hypothesis was that when the experts used the EFRT tool to evaluate the interviewer's skill, scores would differ significantly, and scores on the EFRT would be highest for the good interview, in the mid-range for the fair interview, and lowest for the poor interview. The scores differed significantly and were in the expected range.

To determine intrarater reliability each expert viewed one interview twice, and evaluated the interviewer's skills using the EFRT. The third hypothesis was that for each expert, the scores of the first viewing of each videotaped interview would differ from the scores on the second viewing of the same interview. We rejected hypothesis three for the poor interview and for the fair interview. The scores did not differ significantly between viewings which implies the scores between each viewing were reliable and consistent. We failed to reject hypothesis three for the good videotaped

interview. The scores of the first viewing of the good interview did differ significantly from the scores on the second viewing.

The EFRT scores from the two viewings of the good interview did differ significantly, but this may not be a concern. The good interview sample size was only ten. The sample size was low because only one-third of the original sample of 81 was assigned to view the good video twice. A disproportionate number of the experts who viewed the good video twice either withdrew from the study or did not return two complete EFRT tools. In addition to the small sample size, although the scores did differ significantly, both scores were in the expected range. In a real-world application this variation may be insignificant.

It is possible the methods used for distributing and viewing videotaped interviews may have affected the data. When using the PTA to evaluate the interviews, the videotapes were sent home with the subjects and they had in their possession all three videotaped interviews. All interviews were viewed within a few days of each other, most within 48 hours. In addition, even though the subjects were given instructions as to what order the interviews should be viewed, there was no method of ensuring the subjects followed instructions. In contrast, when the experts used the EFRT, they had only one videotaped interview in their possession. Videotapes were only sent in the order in which they were to be viewed. The experts viewed the interviews with a break of several weeks between each videotaped interview.

Limitations of the Study

The purposive and convenience sample may be a limitation. Subjects could chose to participate, but know little about desirable characteristics of a 24-hour food recall interview. This is probably a minor issue because 74% of subjects reported training other personnel to conduct 24-hour food recall interviews. This could be beneficial because through training they may know what a good interview should include. They may also be accustomed to evaluating interviewers. Their experience may create more accurate evaluations for this study that would generate more accurate data.

Another limitation was to use the fair videotaped interview. The judgmental statement during the interview may have affected how that particular interview was evaluated and made it difficult to discriminate the difference between the fair videotape and the other two videotapes. The mistake could not be edited out of the videotape because without it a person would not understand why the relationship between the interviewer and respondent had changed. After the study, if the experts choose to use the interviews when training their personnel, the fair interview can be used to illustrate how one mistake can alter the interviewer-respondent relationship.

An additional limitation was during the filming of the interviews the interviewer rarely lost focus of the interview. For both the PTA and EFRT, keeping interview on track is a desired characteristic. For all three interviews on both tools the question regarding focus had consistently high scores.

The effect of viewing order was not considered in this study. To compare the scores based on viewing order, the sample would have needed to be much larger. Had it

been possible, it would have been interesting to analyze how scores on both the EFRT and PTA varied with viewing order.

We had two subject populations, but each population only used one evaluation tool. Each subject used only the PTA or EFRT, not both. We did not have the students use the EFRT because when we conducted phase one of the study we needed a quick and easy validation of the videotaped interviews. The experts did not use the PTA because we did not realize how valuable that data could be until after the study was concluded.

CHAPTER VI.

CONCLUSIONS AND RECOMENDATIONS

To increase reliability and to obtain more detailed food recall data the EFRT should include additional items. For this study the primary trait analysis (PTA) measures an interviewer's skill on 17 items, while the EFRT measures only nine items. The reliability of the PTA was slightly higher than the reliability of the EFRT because multiple items on a composite measurement improve reliability (Schutt, 1999). Rather than asking if the interviewer used visual aids, the EFRT should include at least one question about the type of visual aid used. Items should be included to address how food was prepared, the precise names (or brand names) of foods, or if food was commercially prepared. An item should be included to address portion size.

On the EFRT the item, "Was overall food recall accurate?" should be deleted. Unless the interviewer is in training and the respondent has been given a precise list of all foods eaten, the evaluator will have no idea of what is necessary to be considered accurate. Without comparison to a more precise food intake method, no 24-hour food recall interview can be considered accurate (Howat et al, 1994; Olendzki et al, 1999; Tran et al, 2000).

To reduce confusion of the EFRT tool the format should be changed. On the current EFRT most subjects did not answer the yes/no item. The item was confusing because it required the user to circle a choice, rather than check a box. All items should have the same format.

The fair videotaped interview included the negative comment about the respondent's intake of fried food items. For training purposes the fair videotaped interview can be used to show how one mistake can alter the relationship between the interviewer and respondent. Trust and rapport are important to reduce intentional misreporting (IOM, 2000; Windsor et al, 1994 Wolper et al, 1995). Trust can be easily broken by one disapproving statement.

The EFRT was created to be a quick and easy method of evaluating an interviewer's 24-hour food recall skill; it does not require previous training to use. The tool is designed to be used with the *The 24-Hour Recall In-Service Training Program*, but could be used in other situations. The EFRT could be extremely useful in teaching trainees to evaluate interviews. If the evaluator were trained in using the tool, the evaluation would be more reliable and accurate. The tool could be used to train or educate students in college level nutrition assessment courses. The tool would be an acceptable example for developing tools for evaluating other skills. The tool would be very useful in producing numeric data to monitor the effectiveness of a particular interviewer over multiple interviews or the effectiveness of a 24-hour recall interview training program. For personnel managers, the tool could be used for personnel decisions such as performance reviews. Current trends are to conduct interviews over the phone or one paraprofessional simultaneously interviews a group. The EFRT tool could be adapted to evaluate a telephone or group interview, but it should be validated before being used in that capacity (Casey et al, 1999). The multiple pass method of interviewing is considered more accurate than a single pass interview. This form does not address any aspect of the multiple pass interview, and would not be an appropriate evaluation tool.

The EFRT, PTA and videotaped interviews could be used for further research. If a large enough population sample could be acquired, the affect of viewing order on EFRT or PTA score should be investigated. In addition, one population could view the interviews using both the EFRT and PTA. This study could be repeated, with a revised version of the fair interview, omitting the negative statement about the respondent's intake. The tool should be validated for either group or telephone interviews. The tool could be adjusted and validated for multiple pass interviews.

The purpose of this study was to validate a tool to used to evaluate a paraprofessional's ability to conduct an effective 24-hour recall interview. Although the reliability is somewhat limited, the Evaluation of Food Recall Techniques is valid. It should only be used to measure the ability of a paraprofessional to perform a face-to-face 24-hour food recall interview for one individual.

References

Baranowski T, Domel SB. A cognitive model of children's reporting of food intake. *Am J Clin Nutr.* 1994; 59:212S-219S.

Barrett-Conner E. Nutritional epidemiology: how do we know what they ate? *Amer J Clin Nutr.* 1991; 54:182s-187s

Buzzard MI. Rationale for an international conference series on dietary assessment methods. *Amer J Clin Nutr.* 1994; 59:143S-146S.

Casey PH, Goolsby SL, Lensing SY, Perloff BP, Bogle ML. The use of telephone interview methodology to obtain 24-hr dietary recalls. J Amer Diet Assoc. 1999; 99:1406-1412

Frank GC, Berenson GS, Schilling PE, Moore MC. Adapting the 24-hr. recall for epidemiologic studies of school children. J Am Diet Assoc. 1977; 26-31.

Howat PM, Mohan R, Champagne C, Monlezun C, Wozinak P, Bray GA. Validity and reliability of reported dietary intake data. *J Amer Diet Assoc.* 1994; 94:169-174.

Institute of Medicine. *Dietary Reference Intakes: Applications in Dietary Assessment* 2000. National Academy Press, Washington, D.C.

Johnson R, Soultanakis RP, Matthews DE. Literacy and body fatness are associated with underreporting of energy intake in US low-income women using the multiple-pass 24-hour recall: A doubly labeled water study. *J Amer Diet Assoc.* 1998; 98:1136-40.

Jonnalagadda SS, Mitchell DC, Smiciklas-Wright H, Meaker K, Van Heel N, Karmally W, Ershow AG, Kris-Etherton P. Accuracy of energy intake data estimated by a multiple-pass, 24 hour dietary recall technique. J Amer Diet Assoc. 2000; 100:303-309.

Karvetti RL, Knuts LR. Validity of the 24-hour dietary recall. J Amer Diet Assoc. 1985; 85:1437-1442.

Kohlmeier L. Gaps in dietary assessment methodology: meal- vs. list- based methods. *Am J Clin Nutr* 1994;59:175S-180S.

Lee, RD, Nieman, DC. *Nutritional Assessment, 2nd Ed.* 1996. McGraw-Hill, Boston.

McNutt S, Berlin M, Meader J. 1998. Continuing survey of food intake by individuals methodology data collection: training and monitoring interviewers. In: Champagne C, ed. *21st National nutrient databank conference proceedings Champagne, C, ed.* 1996. New Orleans, LA.

Mela DJ, Aaron JI. Honest but invalid: what subjects say about recording their food intake. *J Am Diet Assoc.* 1997; 97:791-793.

Olendzki B, Hurley TG, Herbert JR, Ellis S, Merriam PA, Luippold R, Rider L, Ockene IS. Comparing food intake using the Dietary Risk Assessment with multiple 24-hr recalls and the 7-day Dietary Recall. J Amer Diet Assoc. 1999; 99:14331439.

Schutt RK. *Investigating the social world: the process and practice of research2nd ed.* 1999. Pine Forge Press: Thousand Oaks, California.

Thompson FE, Byers T. Dietary assessment resource manual. *J Nutr.* 1994; 124:2245S-2318S.

Tran KM, Johnson RK, Soultanakis R, Matthews DE. In-person vs telephone-administered multiple-pass 24-hour recalls in women: validation with doubly labeled water. J Amer Diet Assoc. 2000; 100:777-781.

Van Staveren WA, DeGroot L, Blauw YH, Van der Wielen RPJ. Assessing diets of elderly people: problems and approaches. *Am J Clin Nutr.* 1994; 59:221S-224S.

Vogt WP. *Dictionary of statistics and methodology: a nontechnical guide for the social sciences 2nd ed.* 1999. Sage Publications, Inc. Thousand Oaks California.

Walvoord BE, Anderson VJ. *Effective Grading: a tool for learning and assessment.* 1998. Jossey-Bates, Inc. San Francisco, CA.

Williams G. The 24-hour food recall: an essential tool in nutrition education; in-service training program. Oklahoma Cooperative Extension Service, Stillwater, OK 1997.

Windsor R, Baranowski T, Clark N, Cutter G. *Evaluation of health promotion, health education and disease prevention programs, 2nd ed.* 1994 Mayfield Publishing Company Mountain View, CA.

Wolper C, Heshka S, Heymsfield SB. 1995. Measuring food intake: an overview. In: Handbook of Assesment Methods for Eating Behaviors and Weight Related Problems: Measures, Theory, and Research. DB Allison, ed. Sage Publications. Thousand Oaks, CA.

Appendix A

Institutional Review Board Approval Form

Oklahoma State University
Institutional Review Board

Protocol Expires: 6/6/02

Date: Wednesday, October 03, 2001 IRB Application No HE0166

Proposal Title VALIDATION OF AN INSTRUMENT TO MEASURE PARAPROFESSIONALS' ABILITY TO CONDUCT A 24-HOUR FOOD RECALL INTERVIEW

Principal Investigator(s)

Krista Hamilton
414 W 3rd #4
Stillwater OK 74075

Kathryn Kelm
421 HES
Stillwater, OK 74078

Reviewed and Processed as: Expedited

Approval Status Recommended by Reviewer(s): Approved

Dear PI

Your IRB application referenced above has been approved for one calendar year. Please make note of the expiration date indicated above. It is the judgment of the reviewers that the rights and welfare of individuals who may be asked to participate in this study will be respected, and that the research will be conducted in a manner consistent with the IRB requirements as outlined in section 45 CFR 46.

As Principal Investigator, it is your responsibility to do the following:

1. Conduct this study exactly as it has been approved. Any modifications to the research protocol must be submitted with the appropriate signatures for IRB approval.
2. Submit a request for continuation if the study extends beyond the approval period of one calendar year. This continuation must receive IRB review and approval before the research can continue.
3. Report any adverse events to the IRB Chair promptly. Adverse events are those which are unanticipated and impact the subjects during the course of this research; and
4. Notify the IRB office in writing when your research project is complete.

Please note that approved projects are subject to monitoring by the IRB. If you have questions about the IRB procedures or need any assistance from the Board, please contact Sharon Bacher, the Executive Secretary to the IRB, in 203 Whitehurst (phone: 405-744-5700, sbacher@okstate.edu).

Sincerely,

Carol Olson, Chair
Institutional Review Board

Appendix B

Institutional Review Board Continuation Approval Form

Oklahoma State University
Institutional Review Board

Protocol Expires: 5/15/03

Date : Thursday, May 16, 2002 IRB Application No: HE0160

Proposal Title: VALIDATION OF AN INSTRUMENT TO MEASURE PARAPROFESSIONALS' ABILITY TO CONDUCT A 24-HOUR FOOD RECALL INTERVIEW

Principal Investigator(s)

Krista Hamilton
2935 E. 94 Ct #814
Tulsa, OK 74137

Kathryn Keim
421 HES
Stillwater, OK 74078

Reviewed and Processed as: Expedited Continuation

Approval Status Recommended by Reviewer(s) : Approved

Signature : Carol Olson

Carol Olson, Director of University Research Compliance

Thursday, May 16, 2002
Date

Approvals are valid for one calendar year, after which time a request for continuation must be submitted. Any modifications to the research project approved by the IRB must be submitted for approval with the advisor's signature. The IRB office MUST be notified in writing when a project is complete. Approved projects are subject to monitoring by the IRB. Expedited and exempt projects may be reviewed by the full Institutional Review Board.

Appendix C

Primary Trait Analysis of Interviewer Characteristics

Primary Trait Analysis of Interviewer Characteristics

Covers 24 Hours	Yes			No
Interviewer introduces self.				
Interviewer explained purpose of interview.				
Interviewer explained the importance of reporting actual intake.				
Interviewer Characteristics	Usually Performs	Occasionally Performs	Rarely Performs	Does Not Perform
Interviewer does not verbally or nonverbally indicate any approval or disapproval of the respondent's answer.				
Interviewer guides the respondent through the interview using open-ended questions, without suggesting responses.				
Interviewer uses verbal questions and visual cues to help respondent remember what foods were eaten.				
Interviewer asks about activities performed during the day to improve memory of foods eaten.				
Interviewer obtains estimates of the portion size of foods that were eaten.				
Interviewer uses two-dimensional visual aids to help the respondent visualize the type of food eaten and portion size consumed. The models should realistically represent food or portion sizes.				
Interviewer uses three-dimensional visual aids to help the respondent visualize the type of food and portion sizes consumed. The models should realistically represent food or portion sizes.				

Continued other side.

Primary Trait Analysis of Interviewer Characteristics

Interviewer uses measuring cups, measuring spoons, common size plates, bowls or drinking glasses to help the respondent estimate portion size.				
Interviewer asks the specific name of foods that were eaten, and clarifies that the interviewer and respondent are visualizing/conceptualizing the same food.				
Interviewer obtains brand names of foods that were not prepared from a home recipe, including names of restaurants where foods might have been purchased.				
Interviewer asks how foods were prepared and asks what cooking methods (bake, fry, raw) were used.				
Interviewer asks what was added to foods before eating.				
Interviewer keeps interview focused and on track.				
Interviewer avoids meal labeling unless done by the respondent.				

Appendix D

Interview Script Given to Actress

Poor Interview	
Interviewer	Respondent
Hi. I'm going to ask you about what you ate yesterday.	Ok.
What did you eat for breakfast yesterday?	*Gives answer, Interviewer quickly writes down answer on pad in lap.*
Asks one probing question about answer given.	*Gives answer, Interviewer writes response on pad.*
Did you eat anything between breakfast and lunch?	*Gives one word answer.*
What did you eat?	Give brief, one or two word answer. Interviewer writes down response.
You ate fruit for a snack? That's a great snack! Good job!	
Where did you go for lunch?	*Gives answer, doesn't bother to correct assumption if necessary.*
Gives shocked look at mention of restaurant name. Tells story about restaurant interviewee mentions.	
After a few minutes asks about what was eaten	*Gives answer, Interviewer quickly writes down answer.*
Interviewer does not ask specific questions about "lunch".	
What did you have for dinner?	*Gives brief answer, Interviewer writes down answer.*
Reaches into bag near feet, pulls out poor food model. Was the food you ate about this size or was it bigger?	Gives answer, Interviewer writes down.
You added something to your food at the table didn't you?	Gives answer, Interviewer quickly writes down answer.
After dinner did you eat popcorn or chips or anything while you watched tv?	*Gives answer, Interviewer quickly writes down answer.*
You know how many calories are in a bag of microwave popcorn?! Especially if it's that "movie theater butter" kind; that stuff will kill you.	
Ok, did we forget anything? Did you remember all the drinks and candy and gum you ate during the day?	I think so.
Good. Then we're done here. It was nice to meet you. See you next time.	

Fair Interview	
Interviewer	Respondent
Hi. My name is Krista, and I'm going to ask you about what you ate yesterday. It's really important that tell me what you ate.	Ok.
What was the first thing you ate when you got up?	*Gives answer, Interviewer writes down answer on pad in lap.*
Asks one probing question about answer given.	*Gives answer, Interviewer writes response on pad.*
Did you drink anything with that?	*Gives one word answer.*
What did you drink?	Gives one or two word answer. Interviewer writes down response.
What did you do next? Did you go to work, get in the shower? What did you do?	*Gives answer, Interviewer writes response on pad.*
Did you eat or drink anything while you were doing that?	Gives answer, Interviewer writes response on pad.
Shows approval about whatever answer interviewee gives.	
Where was the next thing you did?	*I had lunch at a restaurant. .*
What restaurant, and what did you have?	*Gives answer, Interviewer writes response on pad.*
Gives shocked look at mention of restaurant name. Tells story about restaurant interviewee mentions.	
After a few minutes asks specific questions about what was eaten	*Gives answer, Interviewer quickly writes down answer.*
Interviewer pulls from under pad a paper with poor two dimensional pictures of foods. Was your food bigger than this, smaller than this, about the same size?	*Gives answer, Interviewer writes response on pad.*
What did you have for dinner?	*Gives brief answer, Interviewer writes down answer.*
Reaches into bag near feet, pulls out good food model. Was the food you ate about this size or was it bigger or smaller?	Gives answer, Interviewer writes down.
You added salt to your food at the table didn't you?	Gives answer, Interviewer quickly writes down answer.
After dinner did you eat popcorn or chips or anything while you watched TV?	*Gives answer, Interviewer quickly writes down answer.*

Fair Interview Continued	
You know how many calories are in a bag of microwave popcorn?! Especially if it's that "movie theater butter" kind; that stuff will kill you.	
Ok, did we forget anything? Did you remember all the drinks and candy and gum you ate during the day?	I think so.
What if anything did you eat during the night?	*Gives answer. Interviewer writes response on pad.*
Did you get up at the same time this morning that you got up yesterday?	I don't understand the question.
Yesterday you said you got up about 10:30, did you get up at 10:30 today?	Yes I did.
Good. Then we're done here. It was nice to meet you. See you next time.	

Good Interview	
Interviewer	Respondent
Hi. My name is Krista, and I'm going to ask you about what you ate yesterday. It's really important that tell me what you ate. The information I collect today will be compared to the information I collect in another interview. I'll do the other interview after you attend nutrition classes. I don't need this information to look at what you're eating, but to look at how things change after nutrition education. This interview is what we call "baseline" information.	Ok.
What was the first thing you ate or drank when you got up?	*Gives answer, Interviewer writes down answer on pad in lap.*
Asks at least one probing question about each item mentioned.	*Gives answer, Interviewer writes response on pad.*
What did you do next?	*Gives answer.*
What, if anything, did you eat or drink while you were doing that?	Gives answer, Interviewer writes response on pad.
Where was the next thing you did?	*I had lunch at a restaurant. .*
What restaurant, and what did you have?	*Gives answer, Interviewer writes response on pad.*
Asks at least one probing question about each item mentioned.	
Interviewer pulls from bag near feet the good 2D pictures of foods Which one of these pictures looks closest to the size of the food you ate?	*Points to one of the pictures.*
Was your food bigger than this, smaller than this, about the same size?	*Gives answer.*
Could you show me with your hands about how big it was?	*Demonstrates a circle with hands. Interviewer writes this down.*
What did you do next?	*Gives answer.*
What, if anything, did you eat or drink while you were doing that?	Gives answer, Interviewer writes response on pad.
Asks at least one probing question about each item mentioned	*Answers each question, Interviewer writes all responses on pad.*
Reaches into bag near feet, pulls out good food model. Was the food you ate about this size or was it bigger or smaller?	Gives answer, Interviewer writes down.
What was the next thing you did?	*Gives answer.*

Good Interview Continued	
What, if anything, did you eat or drink while you were doing that?	*Gives answer, Interviewer quickly writes down answer.*
What , if anything, did add to your food at the table?	Gives answer, Interviewer quickly writes down answer.
What was the next thing you did?	*Gives answer.*
What, if anything, did you eat or drink while you were doing that?	*Gives answer, Interviewer quickly writes down answer.*
What if anything did you eat during the night?	*Gives answer, Interviewer writes response on pad.*
Did you get up at the same time this morning that you got up yesterday?	I don't understand the question.
Yesterday you said you got up about 10:30, did you get up at 10:30 today?	Yes I did.
Ok, did we forget anything? Can you think of anything we might have forgotten during the day?	I can't think of any thing.
Then we're done here. It was nice to meet you. See you next time.	

Appendix E

Documentation for Written Informed Consent: Validation of Videotapes

DOCUMENTATION FOR WRITTEN INFORMED CONSENT

Validation of Videotapes

I,_____, hereby authorize or direct Dr. Kathryn S. Keim and Krista Hamilton, or associates their choosing, to perform the following treatment or procedure.

DESCRIPTION OF RESEARCH

This study is to validate videotapes to be used for *Validation of an Instrument to Measure Paraprofessionals' Ability to Conduct a 24-hour Food Recall.* The study involves research and is being conducted through Oklahoma State University Department of Nutritional Sciences and Cooperative Extension Service. The principle investigator is Krista Hamilton, graduate student in the Department of Nutritional Sciences. The purpose of the validation study is to evaluate videotaped interviews.

Procedure

1. View 10-15 minute segments of three videos of 24-hour food recall interviews.

2. Evaluate the videotaped interviewer's skills using the provided primary trait analysis tool.

3. Return all materials to investigators.

Duration of subject's participation

To view and evaluate all three videos will take 30-60 minutes.

Confidentiality of records

Subjects' names will only be written on this consent form. No names will appear on any forms or in any reports

Possible discomforts or risks

There are no anticipated discomforts or risks due to participation in the present study.

Possible benefits for subjects/society

By conducting this study we will collect data to help Nutrition Education Assistants perform more effective 24-hour food recall interviews. By improving interview techniques, food intake data and program evaluation data will be more accurate and valid.

I understand that participation is voluntary and that I will not be penalized if I choose not to participate. I also understand that I am free to withdraw my consent and end my participation in this project at any time without penalty after I notify the project director.

For further information regarding:

Research: Krista Hamilton, Department of Nutritional Sciences, 425 HES, Oklahoma State University, Stillwater, OK, 74078. Phone: 405-624-8296 e-mail: hamiltk@okstate.edu

Research subjects' rights and additional information: Institutional Review Board, Sharon Bacher, IRB Executive Secretary, Oklahoma State University, 203 Whitehurst, Stillwater, OK 74078. Phone: 405-744-5700

I have read and fully understand the consent form. I sign it freely and voluntarily.

Date:_____ Time:_____(a.m./p.m)

Signed:_____

Signature of Subject

I certify that I have personally explained all elements of this form to the subject before requesting the subject to sign it.

Signed: _____

Project director or authorized representative

Appendix F

Instructions to Subjects for Validation of Videotaped Interviews: Version 1

October 1, 2001

Dear Participant:

Thank you for participating in this study. This research is being conducted through Oklahoma State University Department of Nutritional Sciences and Cooperative Extension Service. The purpose of the research is to evaluate videotaped interviews of 24-hour food recalls.

The packet you have received today contains several items. After you have viewed all three videos, please return all items to Dr. Keim.

1) An informed consent form: This details all information regarding confidentiality, possible risks and benefits to society. It also gives information on who to contact regarding this study. Please sign this form and immediately return it to Dr. Keim.
2) Three videos: Each video is labeled with a letter and color.
3) Three evaluation forms: Each form is labeled with a letter and color.

Please view each video in the following order:

Video Tape W—Black Label	Primary Trait Analysis Form W—Grey Paper
Video Tape F—Blue Label	Primary Trait Analysis Form F—Blue Paper
Video Tape Q—Orange Label	Primary Trait Analysis Form Q—Orange Paper

After you view each video, evaluate the interviewer by completing the corresponding form.

After completing the evaluation, please return all materials, including the videotapes within 7 days to Dr. Keim.

If you have questions, please contact Krista Hamilton (918-296-3153, evenings; hamiltk@okstate.edu), or Kathryn S. Keim (405-744-8293, kkathry@okstate.edu).

Thank you for helping.
Sincerely.

Krista Hamilton
Graduate Student
Nutritional Sciences

Kathryn S. Keim
Assistant Professor
Nutritional Sciences

Glenna Williams
Coordinator
CNEP/FCS

Appendix G

Instructions to Subjects for Validation of Videotaped Interviews: Version 2

October 1, 2001

Dear Participant:

Thank you for participating in this study. This research is being conducted through Oklahoma State University Department of Nutritional Sciences and Cooperative Extension Service. The purpose of the research is to evaluate videotaped interviews of 24-hour food recalls.

The packet you have received today contains several items. After you have viewed all three videos, please return all items to Dr. Keim.

1) An informed consent form: This details all information regarding confidentiality, possible risks and benefits to society. It also gives information on who to contact regarding this study. Please sign this form and immediately return it to Dr. Keim.
2) Three videos: Each video is labeled with a letter and color.
3) Three evaluation forms: Each form is labeled with a letter and color.

Please view each video in the following order:

Video Tape F--Blue Label	Primary Trait Analysis Form F—Blue Paper
Video Tape Q—Orange Label	Primary Trait Analysis Form Q—Orange Paper
Video Tape W—Black Label	Primary Trait Analysis Form W—Grey Paper

After you view each video, evaluate the interviewer by completing the corresponding form.

After completing the evaluation, please return all materials, including the videotapes within 7 days to Dr. Keim.

If you have questions, please contact Krista Hamilton (918-296-3153, evenings; hamiltk@okstate.edu), or Kathryn S. Keim (405-744-8293, kkathry@okstate.edu).

Thank you for helping.
Sincerely,

Krista Hamilton	Kathryn S. Keim	Glenna Williams
Graduate Student	Assistant Professor	Coordinator
Nutritional Sciences	Nutritional Sciences	CNEP/FCS

Appendix H

Instructions to Subjects for Validation of Videotaped Interviews: Version 3

October 1, 2001

Dear Participant:

Thank you for participating in this study. This research is being conducted through Oklahoma State University Department of Nutritional Sciences and Cooperative Extension Service. The purpose of the research is to evaluate videotaped interviews of 24-hour food recalls.

The packet you have received today contains several items. After you have viewed all three videos, please return all items to Dr. Keim.

1) An informed consent form: This details all information regarding confidentiality, possible risks and benefits to society. It also gives information on who to contact regarding this study. Please sign this form and immediately return it to Dr. Keim.
2) Three videos: Each video is labeled with a letter and color.
3) Three evaluation forms: Each form is labeled with a letter and color.

Please view each video in the following order:

Video Tape Q—Orange Label	Primary Trait Analysis Form Q—Orange Paper
Video Tape W—Black Label	Primary Trait Analysis Form W—Grey Paper
Video Tape F—Blue Label	Primary Trait Analysis Form F—Blue Paper

After you view each video, evaluate the interviewer by completing the corresponding form.

After completing the evaluation, please return all materials, including the videotapes within 7 days to Dr. Keim.

If you have questions, please contact Krista Hamilton (918-296-3153, evenings; hamiltk@okstate.edu), or Kathryn S. Keim (405-744-8293, kkathry@okstate.edu).

Thank you for helping.
Sincerely,

Krista Hamilton
Graduate Student
Nutritional Sciences

Kathryn S. Keim
Assistant Professor
Nutritional Sciences

Glenna Williams
Coordinator
CNEP/FCS

Appendix I

Recruitment Letter for Validation of Evaluation Tool

«Subject_Name»
«Address»
«Address_2»
«City», «State» «Zip»

April 1, 2001

Dear «Subject_Name»,

We would like to take this opportunity to request your help in a study to be conducted at Oklahoma State University. You were selected because you are a member of the «Division» Division of the Society for Nutrition Education. *The 24-Hour Food Recall In-Service Training Program* was developed by OSU Cooperative Extension Service. The purpose of the program is to train Nutrition Education Assistants (NEA's) to conduct a through 24-hour food recall. The program consists of an instructional video and accompanying booklet. At the end of the booklet is a tool to evaluate the NEA's food recall skills. The purpose of the present study is to validate the NEA 24-hour food recall skill level evaluation tool.

Subjects who volunteer for this study will be asked to view four videotapes of 24-hour food recall interviews and use the evaluation tool to evaluate the interviewers 24-hour recall skills. Each videotape is approximately 60 minutes in length. The tapes will be sent to you one at a time. The minimum time involvement is 12 weeks. Time involvement is dependent on how quickly videotapes are reviewed and returned. Please see attached proposed schedule for more information.

If you choose to participate, the first of four evaluation packets will be mailed to you after you return the enclosed consent form and complete the demographic information form. Each packet will consist of a videotape of an interview and an evaluation tool. If you would like to participate, please return the signed consent form and demographic information form in the enclosed self-addressed stamped envelope.

If you have questions, please contact Krista Hamilton (405-624-8296, evenings; hamiltk@okstate.edu), or Kathryn S. Keim (405-744-8293, kkathry@okstate.edu).

Thank you for helping.
Sincerely,

Krista Hamilton	Kathryn S. Keim	Glenna Williams
Graduate Student	Assistant Professor	Associate Professor
Nutritional Sciences	Nutritional Sciences	Cooperative Extension

Appendix J

Schedule of Subject's Involvement

Schedule of Subject's Involvement

Materials will be sent eight mailing days after materials from previous step have been returned

Step 1	Recruitment packet mailed to subjects
Step 2	Subjects return signed consent and demographic information form
Step 3	Video #1 and evaluation tool sent to subject
Step 4	Video #1 with participant
Step 5	Video #1 returned
Step 6	Video #2 and evaluation tool sent to subjects
Step 7	Video #2 with participant
Step 8	Video #2 returned
Step 9	Video #3 and evaluation tool sent to subjects
Step 10	Video #3 with participant
Step 11	Video #3 returned
Step 12	Video #4 and evaluation tool sent to subjects
Step 13	Video #4 with participant
Step 14	Video #4 returned
Step 15	Any necessary follow-up; send all videos to participants to keep

Appendix K

Demographic Information Form

Subject # _____ **Subject Name**_____
Name will be removed after number is assigned

Demographic Information

1. Your employment status is best described as... Circle one letter.
   a. Full-time
   b. Part-time
   c. Retired
   d. Not employed, **Go to question 4.**

2. Which of the following best describes the settings in which you spend your work time? Please circle all settings where you spend at least 20% of your work time.
   a. Community/public health facility or organization
   b. Acute-care facility
   c. Long-term care facility
   d. Home care agency
   e. School foodservice operation
   f. College/university foodservice
   g. Commercial foodservice operation
   h. Ambulatory/outpatient clinic or office
   i. Pharmaceutical company
   j. Manufacturer/distributor/retailer
   k. College/university faculty
   l. Cooperative extension or other non-formal education
   m. HMO, physician or other healthcare provider
   n. Private practice/primarily individual client counseling
   o. Consultation, primarily to healthcare facilities
   p. Consultation, primarily to other organizations/industries/media
   q. Other (please specify) _____

3. Please circle the job functions which best describe where you spend at least 20% of your work time. Circle all that apply.
   a. Public health/community nutrition
   b. Higher education
   c. Research
   d. Nutrition Education
   e. Clinical services
   f. Foodservices
   g. Public/commercial foodservices
   h. Wellness/disease prevention
   i. Sales/marketing or product development
   j. Personnel training
   k. Nutrition information/communication
   l. Other (please specify)_____

Continued on other side.

4. Please circle the highest degree earned. Circle one letter.
    a. Associates degree(s)
    b. Bachelor's degree(s)
    c. Master's degree(s)
    d. Doctoral degree(s)
    e. Other_____

5. Please select the area(s) that best describe(s) the focus of your continuing professional education. Please circle all that apply.
    a. Community nutrition
    b. Research
    c. Foodservices
    d. Clinical nutrition
    e. Foods/food science
    f. Management
    g. Adult Education
    h. Education
    i. Other

6. Which of the following best describes your job position? Please circle all that apply.
    a. Resident faculty at a university
    b. Cooperative extension faculty at the state level
    c. Cooperative extension faculty at the county level
    d. Does not apply

7. How often do you conduct 24-hour recall interviews? _____
    Please provide the unit of time.

8. Are you responsible for training personnel to do 24-hour food recall interviews? Circle one letter.
    a. Yes
    b. No

9. When the study is over, would you like to have copies of all sample interviews? Circle one letter.
    a. Yes
    b. No

10. For the study would you prefer the interviews in CD or VHS format? Circle one letter.
    a. CD
    b. VHS

Appendix L

Documentation for Written Informed Consent: Validation of Evaluation Tool

DOCUMENTATION FOR WRITTEN INFORMED CONSENT

Validation of Evaluation Tool

I,_____, hereby authorize or direct Dr. Kathryn S. Keim and Krista Hamilton, or associates their choosing, to perform the following treatment or procedure.

Description of research

The name of the project is Validation of an Instrument to Measure Paraprofessionals' Ability to Conduct a 24-hour Food Recall. The study involves research and is being conducted through Oklahoma State University. The principle investigator is Krista Hamilton, graduate student in the Department of Nutritional Sciences. The purpose of the research is to validate a to tool to evaluate 24-hour recall interview techniques.

Procedure

1. View four videos of 24-hour food recall interviews. The subject will have in his or her possession only one video at a time. One video will be viewed twice.

2. Evaluate the videotaped interviewer's skills using the provided evaluation tool. The tool is a one-page form. There are 10 dimensions to evaluate.

3. Return all materials to investigators.

Duration of subject's participation

To view and evaluate each video will take approximately 60 minutes. The full length of the study is dependent on how quickly materials are returned to the investigators. The minimum time involvement is 12 weeks.

Confidentiality of records

Only subject numbers will be written on materials returned from subjects. The key that matches the subject number with the name and address will be kept in a locked file drawer in the investigators office. All tracking will use subject number. No names will appear in any reports.

Possible discomforts or risks

There are no anticipated discomforts or risks due to participation in the present study.

Possible benefits for subjects/society

By conducting this study we will collect data to help Nutrition Education Assistants perform more effective 24-hour food recall interviews. By improving interview techniques, food intake data and program evaluation data will be more accurate and valid.

I understand that participation is voluntary and that I will not be penalized if I choose not to participate. I also understand that I am free to withdraw my consent and end my participation in this project at any time without penalty after I notify the project director.

Continued other side.

For further information regarding:

Research: Krista Hamilton, Department of Nutritional Sciences, 425 HES, Oklahoma State University, Stillwater, OK, 74078. Phone: 405-624-8296 e-mail: hamiltk@okstate.edu

Research subjects' rights and additional contact: Institutional Review Board, Sharon Bacher, IRB Executive Secretary Oklahoma State University, 203 Whitehurst, Stillwater, OK 74078. Phone: 405-744-5700

I have read and fully understand the consent form. I sign it freely and voluntarily.

Date:_____ Time:_____(a.m./p.m)

Signed:_____

Signature of Subject

Appendix M

Order of Viewing Videotaped Interviews by Subject Number

Order of Viewing Videotaped Interviews by Subject Number

Subject Number		A	B	C	D		Tape 4
1		good	fair	poor	4		good
2		fair	poor	4	good		fair
3		poor	4	good	fair		poor
4		4	good	fair	poor		good
5		good	fair	poor	4		fair
6		fair	poor	4	good		poor
7		poor	4	good	fair		good
8		4	good	fair	poor		fair
9		good	fair	poor	4		poor
10		fair	poor	4	good		good
11		poor	4	good	fair		fair
12		4	good	fair	poor		poor
13		good	fair	poor	4		good
14		fair	poor	4	good		fair
15		poor	4	good	fair		poor
16		4	good	fair	poor		good
17		good	fair	poor	4		fair
18		fair	poor	4	good		poor
19		poor	4	good	fair		good
20		4	good	fair	poor		fair
21		good	fair	poor	4		poor
22		fair	poor	4	good		good
23		poor	4	good	fair		fair
24		4	good	fair	poor		poor
25		good	fair	poor	4		good
26		fair	poor	4	good		fair
27		poor	4	good	fair		poor
28		4	good	fair	poor		good
29		good	fair	poor	4		fair
30		fair	poor	4	good		poor
31		poor	4	good	fair		good
32		4	good	fair	poor		fair
33		good	fair	poor	4		poor
34		fair	poor	4	good		good
35		poor	4	good	fair		fair
36		4	good	fair	poor		poor
37		good	fair	poor	4		good
38		fair	poor	4	good		fair

Order of Viewing Videotaped Interviews by Subject Number

39	poor	4	good	fair	poor
40	4	good	fair	poor	good
41	good	fair	poor	4	fair
42	fair	poor	4	good	poor
43	poor	4	good	fair	good
44	4	good	fair	poor	fair
45	good	fair	poor	4	poor
46	fair	poor	4	good	good
47	poor	4	good	fair	fair
48	4	good	fair	poor	poor
49	good	fair	poor	4	good
50	fair	poor	4	good	fair
51	poor	4	good	fair	poor
52	4	good	fair	poor	good
53	good	fair	poor	4	fair
54	fair	poor	4	good	poor
55	poor	4	good	fair	good
56	4	good	fair	poor	fair
57	good	fair	poor	4	poor
58	fair	poor	4	good	good
59	poor	4	good	fair	fair
60	4	good	fair	poor	poor
61	poor	4	good	fair	good
62	fair	poor	4	good	fair
63	good	fair	poor	4	poor
64	4	good	fair	poor	good
65	poor	4	good	fair	fair
66	fair	poor	4	good	poor
67	good	fair	poor	4	good
68	4	good	fair	poor	fair
69	poor	4	good	fair	poor
70	fair	poor	4	good	good
71	good	fair	poor	4	fair
72	4	good	fair	poor	poor
73	poor	4	good	fair	good
74	fair	poor	4	good	fair
75	good	fair	poor	4	poor
76	4	good	fair	poor	good
77	fair	4	good	fair	fair
78	fair	poor	4	good	poor

Order of Viewing Videotaped Interviews by Subject Number

79	good	fair	poor	4	good
80	4	good	fair	poor	fair
81	poor	4	good	fair	poor
82	fair	poor	4	good	good
83	good	fair	poor	4	fair
84	4	good	fair	poor	poor
85	poor	4	good	fair	good

Appendix N

Evaluation of Food Recall Techniques Tool

EVALUATION OF FOOD RECALL TECHNIQUES

COVERAGE Did food Recall cover a 24-hour period?	Yes			No
PREPARATION Did NEA prepare homemaker before Food Recall?(explaining purpose, importance of reporting what was actually eaten, putting homemaker at ease)	Great job!	Good	Fair	Needs improve -ment
APPROVAL/DISAPPROVAL How well did NEA avoid showing approval or disapproval Of food through word of facial expressions during food recall?				
FOCUS Did NEA avoid getting "off track" during interview? If so, was she able to refocus interview?				
USE OF TOOLS/VISUALS Did NEA use proper tools and visuals to confirm serving sizes?				
USE OF QUESTIONS Did NEA ask open-ended questions?				
FOOD CONSUPTION DURING ACTIVITY Did interview questions relate food consumption to activity? Ex. "Did you eat while preparing food? What did you eat while watching TV? Did you eat anything during the night?"				
ADDED INGREDIENTS Did Food Recall include added ingredients (salt, sugar, and fat) and food preparation?				
AVOID "MEAL LABELING" Did NEA avoid labeling meal types (breakfast, lunch, dinner, etc.) unless done so by homemaker?				
COMPLETION OF FORMS Was Food Recall form properly and thoroughly completed?				
ACCURACY Was overall Food Recall interview accurate?				
AVOID "LEADING" QUESTIONS Did NEA avoid asking "leading" questions?				
Total Score:_____	X3	X2	X1	X0
NOTES Open ended question(s) asked by the NEA:				
"Leading" questions asked by the NEA:				
Additional Comments:				

Appendix O

Instruction Letter Sent With Each Videotaped Interview

«Subject_Name»
«Address»
«Address_2»
«City», «State» «Zip»

April 1, 2001

Dear «Subject_Name»,

This is one of four videotaped interviews you have volunteered to evaluate. This study is being conducted through Oklahoma State University Department o f Nutritional Sciences and Cooperative Extension Service. The purpose of the research is to validate a tool used to evaluate nutrition paraprofessionals 24-hour food recall interview skills. Please complete the evaluation tool because your response will be very helpful in validating the tool.

Please view the interview and rate the interviewer using the evaluation tool provided.

After completing the evaluation, please return all materials, including the videotape using the self-addressed, stamped envelope.

If you have questions, please contact Krista Hamilton (405-624-8296, evenings; hamiltk@okstate.edu), or Kathryn S. Keim (405-744-8293, kkathry@okstate.edu).

Thank you for helping.
Sincerely,

Krista Hamilton	Kathryn S. Keim	Glenna Williams
Graduate Student	Assistant Professor	Coordinator
Nutritional Sciences	Nutritional Sciences	CNEP/FCS

Appendix P

Reminder Postcard

JUST A REMINDER...

Last week, a packet containing a videotaped interview and an evaluation tool were mailed to you. If you have already evaluated the interviewer and returned the packet, please accept our sincere thanks. If not, please do so today. We are especially grateful for your help because we believe your response will be very useful in validating the evaluation tool.

If you did not receive a packet or it was misplaced, please call us at 405-744-8293 or e-mail hamiltk@okstate.edu and we will send another one today.

Sincerely,

Krista Hamilton
Graduate Student
Department of Nutritional Sciences
Oklahoma State University

VITA

Krista Kaye Hamilton

Candidate for the Degree of

Master of Science

Thesis: VALIDATION OF AN INSTRUMENT TO MEASURE PARAPROFESSIONALS' ABILITY TO CONDUCT A 24-HOUR FOOD RECALL INTERVIEW

Major Field: Nutritional Sciences

Biographical:

Personal Data: Born in Oxford, Great Britain on July 1, 1974, raised in El Reno, Oklahoma and Stillwater, Oklahoma, married to Brandon Neal, and mother of Kaich Matthew Neal.

Education: Graduated from El Reno High School in May 1992; attended the University of Oklahoma in Norman, Oklahoma; received Bachelor of Science degree in Nutritional Science, Dietetics Major from Oklahoma State University, Stillwater, Oklahoma in June, 1999; completed dietetic internship at Oklahoma State University in 2000; completed the requirements for the Master of Science at Oklahoma State University in December, 2003.

Experience: Registered dietitian since August 2001; community health dietitian for the Oklahoma State Department of Health since December 2000.

Professional Memberships: Oklahoma Dietetic Association, American Dietetic Association (ADA), ADA Pediatric Nutrition Practice Group and ADA Nutrition in Women's Reproductive Health Practice Group