Are We Gathering Reliable Data?
The Need For Measuring Observer Skill in Wildlife Monitoring

A PROFESSIONAL PAPER

by

Ciel Amy Wharton

Submitted to the College of Agriculture and Life Sciences
of Texas A&M University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

May 2006

Wildlife and Fisheries Sciences Department
Wildlife Science

Are We Gathering Reliable Data?
The Need For Measuring Observer Skill in Wildlife Monitoring

A Professional Paper

by

Ciel Amy Wharton

Approved as to style and content by:

Chairman, Advisory Committee- Dr. Jane Packard

Committee Member- Dr. Amanda Stronza

Committee Member- Dr. Fran Gelwick

May 2006

ABSTRACT

Are We Gathering Reliable Data?
The Need For Measuring Observer Skill in Wildlife Monitoring

(May 2006)

Ciel Amy Wharton

The accuracy and reliability of data collected in wildlife surveys can be affected by observer skill. Field sign surveys are especially sensitive to observer effects because of the proficiency required to correctly identify tracks and signs. As a result, there is a need for a method that will systematically measure the skill of participants in wildlife research. CyberTracker Conservation has created one such tool, the Tracker Evaluation. I analyzed the utility of this evaluation system as both a mechanism for assessing observer skill and as a training tool. I present a case study of two Tracker Evaluation workshops for 19 Texas Parks and Wildlife Department employees. All participants improved their scores from the first evaluation workshop (mean = 62%) to the second (mean = 79%) three months later. The mean increase in score was 17 percentage points, with some participants increasing their score by nearly 30 percentage points. In response to an in-house questionnaire, participants stated that the evaluation process measured their tracking skills well (mean = 4 on a Likert Scale of 1 to 5). Participants' level of confidence for correctly identifying animal tracks and sign increased from the first to the second workshop. Overall, participants were very satisfied with the workshops (mean = 5). This case study illustrates that the Tracker Evaluation has the potential to serve as both a local and an international standard for data collectors while simultaneously

functioning as an effective training instrument. With broader application of this system to wildlife research and monitoring programs that use field signs, managers could better understand observer reliability and its implications for interpretation of survey data.

DEDICATION

To my Mom and Dad,

Thanks for being good friends and parents.

ACKNOWLEDGEMENTS

I would like to thank Jane Packard for taking me on as a graduate student in the Ethology Lab and for all her assistance throughout my time at A&M. I am indebted to my committee members, Amanda Stronza and Fran Gelwick, who offered me their time and advice.

I am grateful to Duane Schlitter and John Young who made this project possible by arranging the funding for the certification workshops. Gary Calkins, the Texas Parks and Wildlife Department District 6 Leader, deserves praise for his openness to new ideas and generous support of this project. I am especially thankful for each of the Texas Parks and Wildlife Department staff members that participated in the Tracker Evaluation. They served as exemplary models for enthusiastic learning- maintaining positive attitudes and humor throughout the training experience.

In addition, Mark Elbroch was an invaluable resource and collaborator for this project. I thank him for providing me with helpful feedback on my paper. Louis Liebenberg and CyberTracker Conservation have made a significant contribution to wildlife science worldwide with the creation of the Tracker Evaluation System. This useful tool has also formed the centerpiece of my project, and for this I am grateful. Finally, I couldn't have made it this far without the constant support of Jonah Evans.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

INTRODUCTION

For wildlife managers and conservationists, accurate records of population parameters such as species distribution, richness, and population trends are often necessary for making effective management and policy decisions (Gros et al. 1996; Silveira et al. 2003). Obtaining this information requires efficient and reliable methods for rapid population assessment. Unfortunately, direct sampling (via capture-recapture, radio telemetry, direct observation, etc.) is often costly, time consuming, site-specific, and may interfere with normal animal behavior (Smallwood & Fitzhugh 1991; Beier & Cunningham 1996). As a result, researchers continue to search for reliable ways to monitor the status of target species, which are both inexpensive and non-intrusive (Wilson & Delahay 2001; Sargeant et al. 2003).

Because direct observations are often impractical and cost-prohibitive, researchers have developed several indirect assessment methods. These techniques include documenting animal field sign (tracks, scats, trails, hair, active dens, calls, etc.) via track surveys, scent stations, track plates, scat surveys, and singing bird surveys (Skalski 1991; Gros et al. 1996; Hamm et al. 2003). Indirect survey methods tend to be cost-effective, rapid, and repeatable (Lewison et al. 2001; Gusset & Burgener 2005), and, as a result, sign surveys have become very popular and widely used.

Animal tracking is commonly used in many wildlife surveys (e.g., foxes (*Urocyon cinereoargenteus*): Wood 1959; coyotes (*Canis latrans*): Linhart & Knowlton 1975; cougars (*Puma concolor*): Smallwood & Fitzhugh 1995; otters (*Lutra lutra*): Ruiz-Olmo et al. 2001) and is predominately utilized in presence/absence surveys and transect counts

for determining species richness, distribution, or population trends over time (Silveira et al. 2003). This technique involves determining the presence and activity of an animal by the signs that they leave behind, and is particularly useful for surveying elusive, rare, or nocturnal animals (Stander et al. 1997).

Tracking has been recognized as a valuable tool for studying numerous species throughout many different habitats, yet its reliability has only begun to be assessed (Stander et al. 1997, Zuercher et al. 2003). A noteworthy study by Stander et al. (1997) investigated the scientific merit of tracking by evaluating the skills of four experienced Ju/'Hoan hunters in South Africa. The trackers were tested on their abilities to correctly identify animals by species, age, sex, and distinguish known individuals. In addition, the Ju/'Hoan were asked to interpret different track and sign scenarios depicting predator-prey relationships. Throughout the multi-part study, the trackers illustrated their remarkable tracking identification and interpretation skills. They received a score of 100% for species identification (147 cases), 93% for animal age (69 cases), 97% for animal sex (69 cases), and 94% on known individuals (32 cases). In the test of their ability to interpret scenarios, the Ju/'Hoan scored 99% (interpreting 249 out of 252 cases correct). Overall, they received a score of 98%, correctly identifying and interpreting 557 out of 569 cases. Stander et al.'s study results, therefore, demonstrate that tracking can serve as an accurate and reliable survey tool if the trackers are highly skilled.

The purpose of this paper is to examine observer reliability in wildlife monitoring and consider one option for addressing this source of variation, bias, and error. Through a case study of two evaluation workshops for the Texas Parks and Wildlife Department (TPWD), I assessed the utility of CyberTracker Conservation's Track and Sign

Interpretation Evaluation System as one tool for improving the accuracy and effectiveness of wildlife monitoring. I specifically analyzed the efficacy of this evaluation system as a method for systematically measuring observer skill and as an educational training tool.

As an intern with the Texas Parks and Wildlife Department, under the supervision of Gary Calkins, I helped to organize two evaluation workshops (October 2005, January 2006). Mark Elbroch, a certified senior tracker evaluator (CyberTracker Conservation 2006), conducted both workshops. My methods for collecting information included participant observation, informal conversations, and analysis of data recorded during the workshops. The data were of two types: (1) results from the evaluation workshops and (2) anonymous participant feedback from pre- and post-workshop questionnaires. All procedures followed standard ethical guidelines for the social sciences (Bernard 2000).

In this paper, I first explain the need for a systematic method of measuring observer reliability in wildlife monitoring. Then, I describe one approach for measuring tracking skill that was developed by CyberTracker Conservation. Next, I document how this procedure was used during two evaluation workshops for TPWD employees. Finally, I comment on the suitability of this procedure for documenting and improving skills of participants in other wildlife monitoring and conservation efforts.

THE NEED FOR MEASURING OBSERVER SKILL

Although track surveys are a low-cost and efficient survey method (Beier & Cunningham 1996; Stander 1998; Silveira et al. 2003), the reality of putting this method to use in the field can pose some challenges (Van Dyke et al. 1986). Animal track and sign identification and interpretation is inherently challenging because variations in animal movement, soil type, and weather dramatically affect a track's appearance (Smallwood & Fitzhugh 1989; Stapper 1989). The size and shape of tracks can change depending on the depth and type of substrate, whether it is coarse sand, mud, or fine dust (Kutilek et al. 1983).

While becoming a proficient tracker is an attainable goal, it requires training, practice, and patience (Halfpenny 1986; Vladimirova & Mozgovoy 2003). Ultimately, the effectiveness of monitoring techniques that use track surveys is dependent on appropriate field conditions (i.e., presence of trackable surfaces) and skilled trackers (Silveira et al. 2003; Sadlier et al. 2004).

Although observer skill is important to address in order to ensure the collection of accurate data, the issue has not received adequate consideration (Davis 1981; Wilson & Delahay 2001; Diefenbach et al. 2003). Anderson (2001) acknowledges that the observer (e.g., education and training), the environment (e.g., temperature and precipitation), and characteristics of the target species (e.g., coloration and behavior) are major sources of bias. However, it is not common practice to directly address all of these variables during the design of a wildlife study. In a review of 75 peer-reviewed ornithology journal papers from 1985-2001, 76% provide no information on the number of observers, and

only 3% report actions taken to reduce the effects of observer variability (Diefenbach et al. 2003).

Research results can be significantly impacted by the skill of participating observers because monitoring methods based on the detection of field signs rely upon correct identification (Wilson & Delahay 2001; Sadlier et al. 2004). Even the most experienced and well-trained observers vary considerably and are neither perfect nor identical in their levels of skill and accuracy (Kepler & Scott 1981). As a result, potential sampling error can be high in wildlife surveys, as misinterpretation of tracks may lead to either overestimation or underestimation of the populations being studied (Reid et al. 1987).

Bird survey results have shown that observer differences have reduced population estimate precision and increased sampling variance (Diefenbach et al. 2003). Analysis of the results of the North American Breeding Bird Survey (BBS) shows that underqualified observers (i.e. those with hearing loss or lack of training) consistently record lower species totals than qualified observers on the same routes (Faanes & Bystrak 1981). BBS analysis also indicates that poorly trained observers in many cases record higher numbers of the species that they are familiar with than do well-trained observers on the same routes (Faanes & Bystrak 1981). These findings imply that bias and variability in data due to observer effects may prevent researchers from gaining an accurate understanding of the wildlife populations being studied.

In order to conduct effective wildlife monitoring, it is apparent that observer reliability should be addressed and observers be selected who are highly skilled in correctly identifying target species or their field signs. It is common for project managers

to choose their research staff based on academic background and job-related experience. However, experience is not always correlated with skill level (Kepler & Scott 1981; Freilich & LaRue 1998). Even the reliability of experienced wildlife biologists may vary greatly (Evans 2006). This phenomenon poses a problem for research that depends on skilled observers and illustrates the need for a standardized method to measure the degree of expertise of observers.

A TRACKING SKILL ASSESSMENT TOOL

Testing is commonly used for establishing trust and credibility across professional disciplines from law to medicine. In response to the growing need for skilled wildlife trackers in South Africa, Louis Liebenberg created the Tracker Evaluation System in 1994 (CyberTracker Conservation 2006). This system is a comprehensive method for measuring how accurately observers identify and interpret animal tracks and other field signs. The Tracker Evaluation System has helped local communities by certifying their tracking proficiency. As a result, native Bushman, Shangaan, and Xhosa trackers, as well as rangers and naturalists, have acquired credibility and jobs in wildlife conservation (L. Liebenberg, personal communication). This evaluation system has also been used in South Africa as a training method for National Parks employees (M. Elbroch, personal communication).

Within the Tracker Evaluation System, there are two types of certifications: the Tracker Certification (track/sign identification and trailing) and the Track and Sign Interpretation Certification (only track/sign identification). This paper is focused on the latter, but more information on the entire Tracker Certification process can be found at CyberTracker Conservation's website (http://cybertracker.org).

During a Track and Sign Interpretation Evaluation, participants are tested in the field over two or more days on their track and sign identification and interpretation skills (M. Elbroch, personal communication). Certified evaluators conduct the evaluation in multiple locations within a given area to assess participants' knowledge of a variety of species, tracking substrates, and habitats. Anything from an insect trail to a mammal scat to a bird track can be tested (for examples of questions, see Appendix A). An evaluator

selects a series of tracks and signs encountered in the field and participants are given as much time as needed to independently answer pertinent identification or interpretation questions. Sample questions include: What species made this track or sign? Which foot is this? (Front or hind? Left or right?) How is this animal moving? What was this animal doing?

The system was designed to objectively evaluate the tracking skill of the participant; therefore questions are straightforward and not intended to trick participants (M. Elbroch, personal communication). Questions are discussed in detail by the evaluators after participants have answered a set of questions either verbally or in writing. The scoring system that is used is weighted (Table 1), where the point value of a given question depends upon the species, condition of the track or sign, and the context (CyberTracker Conservation 2006). The difficulty rating of a track is increased when species with similar looking tracks are present (M. Elbroch, personal communication). While detailed track-rating guidelines have been created for many species, the evaluator ultimately determines the difficulty level of each question.

Table 1. Standardized CyberTracker Conservation guidelines for the Tracker Evaluation System.

	Weight of Question (points)		
	Easy (1)	*Difficult (2)*	*Very Difficult (3)*
Track or sign rating criteria	Clear, complete, typical, and/or no similar species	Unclear, incomplete, typical, and/or no similar sp.	Obscure, partial, atypical, and/or similar sp.
Correct	1+	2+	3+
Incorrect	3x	2x	1x

During the Track and Sign Interpretation evaluation, a minimum of 35 questions must be asked with no more than 20% being "easy" or "very difficult" questions. A participant's final score is determined by dividing the sum of all the correct (+) points by the sum of all the correct (+) points and incorrect (x) points, then multiplying by 100 to express the score as a percentage (CyberTracker Conservation 2006). Certificates are issued based on point categories: Track and Sign: Level 1 (Score of 70-79%), Level 2 (Score of 80-89%), and Level 3 (Score of 90-99%). As a result of the weighted scoring system, it is increasingly harder to attain each higher certification level. The participants who receive a score of 100% can choose to take part in a separate "senior" evaluation in which they are asked a minimum of 50 "very difficult" (3-point) questions. Those who get 100% correct are awarded the highest honor, a Track and Sign Specialist Certificate.

Since its creation over ten years ago, the skills of over 400 people have been evaluated using the Tracker Evaluation System (M. Elbroch, personal communication). The system is internationally recognized; it is currently being used in South Africa, Botswana, North America (see: http://wildlifetrackers.com/evals), and will soon be established in Spain (L. Liebenberg, personal communication). By design, this evaluation system includes questions on all species present in the evaluation area, rather than just target species and their look-alikes. This provides participants with a broader, more comprehensive knowledge of animal tracking that helps them distinguish the tracks and sign of their target species in the presence of other species. Using this international standard, performance of participants around the world can be compared. But, how could this system be applied to local or regional monitoring programs in North America?

CASE STUDY: TEXAS PARKS AND WILDLIFE DEPARTMENT TRACK AND SIGN INTERPRETATION WORKSHOPS

BACKGROUND

In the present pilot project, the Texas Parks and Wildlife Department chose CyberTracker Conservation's Tracker Evaluation System as a way to measure observer reliability and explore the system's suitability as a training process for employees in District 6 (G. Calkins, personal communication). In the Piney Woods ecoregion of Texas, the northern river otter (*Lontra canadensis*) has been the target species for track surveys conducted under bridges (Bartnicki & Boone 1989; McGinty & Young 2003; Young 2003).

Otter population trends have been monitored by TPWD because the species is listed in the Convention on the International Trade of Endangered Species (CITES) under Appendix II (CITES 2005). To remain in compliance with the CITES treaty, the U.S. Fish and Wildlife Service requires state agencies to provide evidence that otter populations are harvested in a sustainable manner, in return for issuing federal permits for the export of otter pelts from the state. As a result, TPWD employees conduct "bridge surveys," recording the presence of otter tracks as well as other furbearer species every three years. Surveys are conducted between January and March under approximately 250 bridges in 27 east Texas counties (Evans 2006).

In an effort to improve the TPWD otter-monitoring program, two evaluation workshops were conducted in October 2005 (Workshop I) and January 2006 (Workshop II). Total participants ranged from 20 at the first workshop to 22 at the second, with 19 TPWD employees attending both workshops. Not all of the participants had previous

experience with bridge surveys. The workshops were conducted at a number of bridges, wetland, and upland areas in Jasper and Newton Counties that represent a range of conditions in the otter survey region. The first workshop enabled participants to get a clear idea of their baseline skill level, strengths and weaknesses, and highlighted areas for future improvement and training. The second workshop, three months later, provided an opportunity for participants to demonstrate improvement and to continue to hone their tracking skills.

Although TPWD bridge surveys primarily focus on monitoring river otters, CyberTracker Conservation evaluation procedures were followed during both workshops and the participants were tested on the tracks and sign of a variety of species encountered in the survey region. This experience can help participants determine whether an imprint on the ground was made by an otter or by another animal or object capable of leaving an "otter-like" mark. Ultimately, this fosters a more comprehensive understanding of the challenges inherent in identifying and interpreting tracks and sign.

For example, during Workshop II the participants came upon a particularly challenging question that tested their observation skills. The question was "who or what made these tracks" and pointed to a trail of dots on the ground. Answers ranged from a crayfish to a beetle and a frog, but the mystery required participants to expand their understanding of tracking. While all of the participants spent adequate time examining the trail and thinking, only one person followed the trail down the slope to where the answer laid. Sitting near the water's edge was the spiky seed ball of a sweet gum tree.

In addition to being a method for assessing observer reliability, the evaluation also functions as an educational tool. During both workshops, participants expressed

how useful the evaluation experience was and how much they were learning. One participant explained how learning to identify animal tracks reminded them of learning to see images in a 3-D poster. Paraphrasing this participant, "a track just looked like a mark in the dirt until the evaluator explained it and pointed out the toes, heel pad, etc. Then, the animal print would suddenly jump out and the 3-D image would become crystal clear." This account and the results of the evaluation workshops demonstrate that the Tracker Evaluation process served as an effective training exercise for the TPWD participants, in addition to providing measures of observer skill. An additional source of evidence was the change in the scores of participants, as described below.

WORKSHOP RESULTS

The results of Workshop II demonstrated an overall improvement in tracking skill (Figure 1). Participants' (*n = 19*) scores from Workshop II (mean = 79%, SD = 8%) were substantially higher than Workshop I (mean = 62%, SD = 9%). Participants improved their scores by a mean increase of 17 percentage points. The highest score for Workshop II was 90%, whereas for Workshop I it was 74%. The improvement in participant skill from Workshop I to Workshop II was also made apparent with a comparison of the certificates awarded. At the end of Workshop II, one "Level 3," ten "Level 2," and six "Level 1" certificates were received, whereas only six "Level 1" certificates were awarded at the end of Workshop I.

Score improvement occurred within, as well as between, workshops. On the first day of Workshop I, participants identified an average of 50% of the questions correctly. On the second day, participants averaged 67% correct. Yet, more decisive evidence that

the Tracker Evaluation System can be an effective training tool was how each participant improved their scores between the first and second workshop. For example, with practice and determination, one participant's score increased from 69% in Workshop I to 90% in Workshop II.

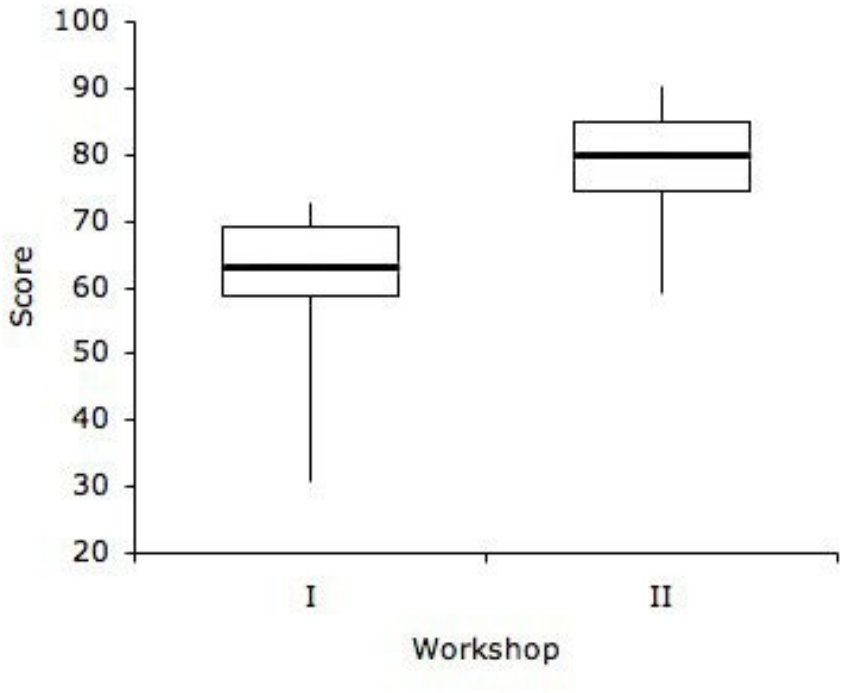

Figure 1. Box plots comparing participant (n = 19) scores from the first to the second TPWD evaluation workshop. Centerlines represent the median, boxes delineate the first and third quartiles, and whiskers delineate the range.

PARTICIPANT FEEDBACK

Each participant was asked to respond to a series of questions on an in-house questionnaire designed in a pre- and post-test format (Appendix B). Most of the feedback questions were closed-ended, with the majority being Likert-scale or multiple-choice questions. Pre- and post-workshop question sets varied slightly, but included

many of the same questions. The post-workshop question set contained additional questions about the effectiveness of the evaluation system, satisfaction level, suggestions for improvement, etc. To compare the two workshops, responses given at the end of the workshops were the focus of this analysis (Table 2).

Table 2. Participant responses to select questions from a questionnaire given at the end of the two Tracker Evaluation workshops.

	Workshop I Mean (n = 19)	Workshop II Mean (n = 19)	Increase	Likert Scale
Level of Confidence-Otter Tracks[a]	3.3	3.8	0.5	1--------2--------3--------4--------5 Very Low Medium Very High
Level of Confidence-Other Tracks[b]	3.0	3.5	0.5	1--------2--------3--------4--------5 Very Low Medium Very High
Skill Measurement[c]	4.2	4.4	0.2	1--------2--------3--------4--------5 Very Poor Satisfactory Very Well
Satisfaction[d]	4.6	4.9	0.3	1--------2--------3--------4--------5 Very Unsatisfied Neutral Very Satisfied
Recommend to others[e]	4.7	4.9	0.2	1--------2--------3--------4--------5 Definitely Not Not Sure Definitely Yes

[a] *What is your current level of confidence for correctly identifying river otter tracks?*
[b] *What is your current level of confidence for correctly identifying other animal tracks and sign?*
[c] *How well did this process measure your tracking skills?*
[d] *How satisfied are you with this workshop?*
[e] *Would you recommend this certification to other TPWD districts?*

Based on participant feedback gathered from the questionnaires (Appendix C), satisfaction with the evaluation workshops was high (Table 2). Confidence scores increased for identifying both otter tracks and other animal tracks. The evaluation procedure received high scores on its ability to measure the tracking skills of participants.

Most participants indicated that they would like evaluation workshops to be offered either every year or every two years, and that they would recommend the procedure to others.

On average, participants reported that they practiced track identification three to four times between workshops. Therefore, caution is recommended when interpreting exactly how much of the learning and improvement could be attributed solely to the participants' experience during the workshops. Based on informal discussions, I believe the success of the workshops could be attributed to several factors, including: (1) the educational elements of CyberTracker Conservation's Evaluation System (i.e. group discussion and evaluator tips for correct identification occurred after participants formulated their own answers), (2) the expertise and communication skills of the workshop facilitator, (3) the participants' personal motivation for self-improvement, and (4) the short time period (3-months) between the two workshops, which enabled information from the first workshop to remain fresh in the participants' minds.

BENEFITS AND LIMITATIONS OF THE TRACKER EVALUATION

Based on the results of the TPWD workshops, I believe the CyberTracker Conservation Track and Sign Interpretation Evaluation System can serve the dual purposes of measuring observer reliability and training participants in track and sign identification and interpretation. Across disciplines, tests have created an incentive for learning. In the field of animal tracking, testing or evaluation establishes a need for learning that can attempt to substitute for the historic driving force behind tracking skill, the need for food (Liebenberg 1990).

CyberTracker Conservation effectively created a tool that could be used by wildlife scientists to address observer reliability as a source of error in wildlife monitoring and research. This evaluation system is an internationally recognized method for measuring the tracking skill in a local context; therefore it can serve as a measure of observer reliability. The Tracker Evaluation has been utilized in South Africa for over ten years, and has provided local trackers with an opportunity to gain recognition for their expertise. In South Africa, trackers generally are hired to conduct wildlife research only if they scored 90% or higher on the Tracker Evaluation (M. Elbroch, personal communication). Because the system has come to be well respected and trusted in South Africa, it has been used as a training technique and has provided certified trackers with the credibility needed to obtain jobs in wildlife research, ecotourism, and anti-poaching.

Despite its apparent utility as a method for measuring tracking skill and as a training tool, CyberTracker Conservation is faced with the challenge of maintaining an international standard for tracking. Yet, this challenge is not unlike those faced by other administrators of certification programs and can be surmounted. As the evaluation

system grows in geographic scope and popularity, it must be continuously monitored to ensure that the standard for trackers maintains consistency and integrity (L. Liebenberg and M. Elbroch, personal communication).

Many precautions have already been taken to preserve the reliability of the system. Prior to becoming an evaluator, candidates must have achieved a score of 100% on the tracker evaluation and are required to serve as an assistant evaluator until they have a clear understanding of the scoring system (CyberTracker Conservation 2006). Also, the system is structured in Africa so that there is an "external" evaluator who is familiar with evaluation standards throughout the broader evaluation region that oversees "local" evaluators in order to maintain uniformity in testing. In addition, the evaluation procedure and scoring system are concisely documented on the CyberTracker website. A set of guidelines for track and sign ratings has been established for species in South Africa, and a similar resource is currently being created for North America (M. Elbroch, personal communication).

Despite these concerted efforts, the evaluation system will inevitably encounter growing pains. While some variation will naturally result from factors like the strengths and weaknesses of different evaluators, inconsistencies should not be allowed to compromise the goal of maintaining an international standard for tracking.

POTENTIAL APPLICATIONS

The future appears to hold promise for certified wildlife trackers, as their skills continue to be needed for wildlife monitoring, research, and ecotourism. By encouraging broad acceptance and use of the Tracker Evaluation System, one goal of CyberTracker Conservation is to set a high, credible standard for wildlife tracking in North America and beyond. The potential applications of the Tracker Evaluation and similar skill assessment techniques are both vast and ultimately limited by specific project needs.

The Tracker Evaluation could be used as a screening tool, where managers or primary investigators select applicants for specific monitoring tasks, based on the tracking certification level in addition to evidence of university and employment performance. Currently tracking is rarely taught within North American university systems, and as a result tracking expertise may not be commonly found even among wildlife professionals. Animal tracking relies on a specialized knowledge base (knowledge of animal life history, ecology, foot morphology, animal movement patterns, etc) and requires intensive training and/or experience (Halfpenny 1986; Silveira et al. 2003). Despite the fact that identification of tracks and sign can be difficult (Smallwood & Fitzhugh 1989; Stapper 1989), 100% correct identification of field sign is possible and highly skilled trackers do exist (Stander et al. 1997; Liebenberg et al. 1999; Zuercher et al. 2003). Unfortunately, financial constraints and academic bias (i.e., hiring only persons who have wildlife degrees) may cause skilled fieldworkers to be overlooked in wildlife research and management (Wilson & Delahay 2001). Therefore, there may be

people in the public sector whose proficient tracking skills are not being utilized. The Tracker Evaluation system could help expand the pool of potential observers.

An additional application of the Tracker Evaluation could be as a quality control measure for "Citizen Science" or volunteer/amateur assisted research. Volunteers constitute a major human resource for conservation organizations, and are invaluable for accomplishing tasks that lack sufficient funding (McLaughlin & Hilts 1999; Foster-Smith & Evans 2003; Newman et al. 2003). Through their participation in a variety of studies, volunteers have contributed greatly to scientific knowledge (Foster-Smith & Evans 2003). Despite the apparent advantages of using volunteers in wildlife conservation, there remains skepticism within parts of the scientific community about the reliability of amateur collected data (Foster-Smith & Evans 2003). For monitoring programs that conduct field sign surveys, the Track and Sign Interpretation Evaluation could help to address these concerns in a systematic manner. As a result, wildlife managers could utilize the skills of volunteers and professional researchers more effectively by assigning them tasks that are appropriate to their certification level (Stander 1998).

CONSERVATION IMPLICATIONS

Within the field of conservation biology, a considerable amount of time, money, and energy is spent on estimating population trends where a decline would signify a population that might enter into threatened or endangered status unless prevented by human intervention (MacKenzie & Kendall 2002). Many population studies rely on indirect indices, such as animal tracks and scats. Therefore, study results often depend on the abilities of surveyors to correctly identify or interpret the field signs in question.

Becoming proficient at track and sign identification/interpretation requires training and practice; therefore it is reasonable to ask about the potential bias of wildlife research that does not explicitly report measures of observer skill. Despite this need to clarify bias, few studies have been conducted on the effect of observers on wildlife monitoring (Verner & Milne 1990). Instead, it is commonly assumed that professional biologists or experienced observers will all score high on reliability measures. However, poor management decisions leading to continued species declines is an expensive price to pay for acting upon ungrounded assumptions.

There is general agreement among the scientific community that management and conservation practices should be founded on sound science. Unfortunately, time and budget constraints can result in hasty decisions based on incomplete knowledge and limited data (Karanth et al. 2003). This reality forces us to ask the question: is an unreliable estimate better than none? Some wildlife researchers state that it is not and that erroneous data actually hinders progress toward possible solutions (Romesburg 1981; Karanth 1999; Anderson 2001). Therefore, to effectively monitor wildlife populations

and make good management decisions, concerted efforts should be made to reduce error and increase data reliability and accuracy. Observers are an indisputable source of variation and bias that should be addressed, for there are striking differences in the identification and interpretation skills of observers - even among those who are highly trained.

Field sign surveys are routinely conducted to study wildlife, yet the observers' abilities to correctly identify the tracks or sign in question are not regularly evaluated. Without knowing the reliability of data collectors, there is a considerable gap in our knowledge and the accuracy of survey results remains difficult to interpret. Thus, the need for observer skill assessment is great. CyberTracker Conservation has created one possible solution, the Tracker Evaluation System. This standardized method for measuring an observer's tracking skill has enabled competent trackers in South Africa to prove their credibility, obtain employment, and gain respect for their knowledge. In addition to assessing observer skill, the TPWD case study illustrated how the system can also function as an effective training tool.

As wildlife researchers continue to move away from single-species management and toward broader scale ecosystem or eventually global monitoring programs, the Tracker Evaluation could prove useful by uniformly and objectively comparing observers and maintaining a high standard for data collectors worldwide. In a world characterized by ecological unpredictability, this tool can help identify the true field sign experts and thereby reduce the variability caused by observers. Broad-scale use of this or similar assessment systems could help researchers determine whether they are detecting real changes in wildlife populations or if observer skill levels are restricting inferences that

could be made from field sign surveys. With the help of a network of certified trackers, perhaps wildlife researchers will be able to more accurately monitor wildlife and managers will be able to make more informed and effective decisions.

LITERATURE CITED

Anderson, D. R. 2001. The need to get the basics right in wildlife field studies. Wildlife Society Bulletin **29**:1294-1297.

Bartnicki, P. L., and D. B. Boone. 1989. Special report: Status of the river otter (*Lutra lutra*) in Texas. Texas Parks and Wildlife Department, Jasper, Texas.

Beier, P., and S. C. Cunningham. 1996. Power of track surveys to detect changes in cougar populations. Wildlife Society Bulletin **24**:540-546.

Bernard, H. R. 2000. Social research methods: qualitative and quantitative approaches. Sage Publications, Thousand Oaks, California.

[CITES] Convention on International Trade in Endangered Species of Wild Flora and Fauna. 2005 June 23. CITES Appendices I, II, and III. Available from http://www.cites.org/eng/app/appendices.html (accessed March 2006).

CyberTracker Conservation. 2006. CyberTracker Conservation home page. Available from http://cybertracker.org/ (accessed March 2006).

Davis, D. E. 1981. Summarizing remarks: observer variability. Studies in Avian Biology **6**:390.

Diefenbach, D. R., D. W. Brauning, and J. A. Mattice. 2003. Variability in grassland bird counts related to observer differences and species detection rates. The Auk **120**:1168-1179.

Evans, J. W. 2006. Observer error in identifying species using indirect signs: analysis of a river otter track survey technique. M.S. thesis. Texas A&M University, College Station, Texas.

Faanes, C. A., and D. Bystrak. 1981. The role of observer bias in the North American Breeding Bird Survey. Studies in Avian Biology **6**:353-359.

Foster-Smith, J., and S. M. Evans. 2003. The value of marine ecological data collected by volunteers. Biological Conservation **113**:199-213.

Freilich, J. E., and E. L. LaRue. 1998. Importance of observer experience in finding desert tortoises. Journal of Wildlife Management **62**:590-596.

Gros, P. M., M. J. Kelly, and T. M. Caro. 1996. Estimating carnivore densities for conservation purposes: indirect methods compared to baseline demographic data. Oikos **77**:197-206.

Gusset, M., and N. Burgener. 2005. Estimating larger carnivore numbers from track counts and measurements. African Journal of Ecology **43**:320-324.

Halfpenny, J. 1986. A field guide to mammal tracking in North America. Second edition. Johnson Publishing Company, Boulder, Colorado.

Hamm, K. A., L. V. Diller, and R. R. Klug. 2003. Spatial independence of Fisher (*Martes pennanti*) detections at track plates in northwestern California. The American Midland Naturalist **149**:201-210.

Karanth, K. U. 1999. Counting tigers with confidence. In Riding the tiger: tiger conservation in human-dominated landscapes:350-353. Seidensticker, J., S. Christie, and P. Jackson (Eds). Cambridge: Cambridge University Press.

Karanth, K. U., J. D. Nichols, J. Seidensticker, E. Dinerstein, J. L. David Smith, C. McDougal, A. J. T. Johnsingh, R. S. Chundawat, and V. Thapar. 2003. Science deficiency in conservation practice: the monitoring of tiger populations in India. Animal Conservation **6**:141-146.

Kepler, C. B., and J. M. Scott. 1981. Reducing bird count variability by training observers. Studies in Avian Biology **6**:366-371.

Kutilek, M. J., R. A. Hopkins, E. W. Clinite, and T. E. Smith. 1983. Monitoring population trends of large carnivores using track transects. Pages 104-106. Proceeding from an International Conference, August 15-19, 1983, Corvallis, Oregon.

Lewison, R., E. L. Fitzhugh, and S. P. Galentine. 2001. Validation of a rigorous track classification technique: identifying individual mountain lions. Biological Conservation **99**:313-321.

Liebenberg, L. 1990. The art of tracking: the origin of science. New Africa Books, Cape Town, South Africa.

Liebenberg, L., L. Steventon, K. Benadie, and J. Minye. 1999. Rhino tracking in the Karoo National Park. Pachyderm, Jan-Dec, Number 27.

Linhart, S. B., and F. F. Knowlton. 1975. Determining the relative abundance of coyotes by scent station lines. Wildlife Society Bulletin **3**:119-124.

MacKenzie, D. I., and W. L. Kendall. 2002. How should detection probability be incorporated into estimates of relative abundance? Ecology **83**:2387-2393.

McGinty, K. D., and J. Young. 2003. Fur-bearing animal recommendations. Federal Aid in Wildlife Restoration Project W-134-2. Small Game and Furbearer Research and Surveys. Texas Parks and Wildlife Department. Austin, Texas.

McLaughlin, L., and S. Hilts. 1999. Monitoring accuracy and the decomposition of error committed by volunteers in a wetland wildlife monitoring program. Leading Edge

Conference Proceedings, October, 1999. Ontario's Niagara Escarpment, Ontario, Canada.

Newman, C., C. D. Buesching, D. W. Macdonald. 2003. Validating mammal monitoring methods and assessing the performance of volunteers in wildlife conservation. Biological Conservation **113**:189-197.

Reid, D. G., M. B. Bayer, T. E. Code, and B. McLean. 1987. A possible method for estimating river otter, *Lutra canadensis*, populations using snow tracks. The Canadian Field-Naturalist **101**:576-580.

Romesburg, H. C. 1981. Wildlife science: gaining reliable knowledge. Journal of Wildlife Management **45**:293-313.

Ruiz-Olmo, J., D. Saavedra, and J. Jimenez. 2001. Testing the surveys and visual and track censuses of Eurasian otters (*Lutra lutra*). Journal of Zoology **253**:359-369.

Sadlier, L. M. J., C. C. Webbon, P. J. Baker, and S. Harris. 2004. Methods of monitoring red foxes *Vulpes vulpes* and badgers *Meles meles*: are field signs the answer? Mammal Review **34**:75-98.

Sargeant, G. A., D. H. Johnson, and W. E. Berg. 2003. Sampling designs for carnivore scent-station surveys. Journal of Wildlife Management **67**:289-296.

Silveira, L., A. T. A. Jácomo, and J. A. F. Diniz-Filho. 2003. Camera trap, line transect census, and track surveys: a comparative evaluation. Biological Conservation **114**:351-355.

Skalski, J. R. 1991. Using sign counts to quantify animal abundance. Journal of Wildlife Management **55**:705-715.

Smallwood, K. S., and E. L. Fitzhugh. 1989. Differentiating mountain lion and dog tracks. Pages 58-63 in R. H. Smith, editor. Proceedings of the Third Mountain Lion Workshop, Dec. 6-

8, 1988. Arizona Game and Fish Department, Prescott, Arizona.

Smallwood, K. S., and E. L. Fitzhugh. 1991. The use of track counts for mountain lion population census. Pages 59-67 in C. E. Braun, editor. Proceedings from Mountain Lion-Human Interaction Symposium. Colorado Division of Wildlife, Denver, Colorado.

Smallwood, K. S., and E. L. Fitzhugh. 1995. A track count for estimating mountain lion *Felis concolor californica* population trend. Biological Conservation **71**:251-259.

Stander, P. E. 1998. Spoor counts as indices of large carnivore populations: the relationship between spoor frequency, sampling effort and true density. Journal of Applied Ecology **35**:378-385.

Stander, P. E., Ghau, D. Tsisaba, and Ui. 1997. Tracking and the interpretation of spoor: a scientifically sound method in ecology. Journal of Zoology **242**:329-341.

Stapper, R. J. 1989. A comparison of scent-station surveys and track counts for surveying furbearer populations in the Big Thicket National Preserve, Texas. M.S. thesis. Texas A&M University, College Station, Texas.

Van Dyke, F. G., R. H. Brocke, and H. G. Shaw. 1986. Use of road track counts as indices of mountain lion presence. Journal of Wildlife Management **50**:102-109.

Verner, J., and K. A. Milne. 1990. Analyst and observer variability in density estimates from spot mapping. The Condor **92**:313-325.

Vladimirova, E., and J. Mozgovoy. 2003. Sign field theory and tracking techniques used in studies of small carnivorous mammals. Evolution and Cognition **9**:1-17.

Wilson, G. J., and R. J. Delahay. 2001. A review of methods to estimate the abundance of terrestrial carnivores using field signs and observation. Wildlife Research **28**:151-164.

Wood, J. E. 1959. Relative estimates of fox population levels. Journal of Wildlife Management **23**:53-63.

Young, J. 2003. Performance report, as required by U.S. Fish and Wildlife Service Division of Management Authority: River otter status. Texas Parks and Wildlife Department, Austin, Texas.

Zuercher, G. L., P. S. Gipson, and G. C. Stewart. 2003. Identification of carnivore feces by local peoples and molecular analyses. Wildlife Society Bulletin **31**:961-970.

APPENDIX A

SAMPLE EVALUATION QUESTIONS

EASY (1-POINT) QUESTIONS

Alligator Trail

Turtle Trail
(Trail direction was a 2-pt. question)

Raccoon Tracks
(Which foot was a 2-pt. question)

DIFFICULT (2-POINT) QUESTIONS

River Otter Tracks

Nutria Tracks

Coyote Track
(Animal movement speed was a 3-pt. question)

VERY DIFFICULT (3-POINT) QUESTIONS

Woodrat Tracks

Great Egret Track

River Otter Scat

APPENDIX B

QUESTIONNAIRES

PRE-WORKSHOP
October 31, 2006

The purpose of these questions is not to evaluate you; rather it is for me to evaluate the effectiveness of the certification process. This is a completely anonymous exercise, and your name will not be used in any way.
Thanks very much for your time!
-Ciel Wharton, Texas A&M, Wildlife and Fisheries Department

<Circle Your Answers>

1. **What is your current level of confidence for correctly identifying river otter tracks?**

   1----------2----------3----------4----------5
   Very Low Medium Very High

2. **What is your level of confidence for identifying other animal tracks and sign?**

   1----------2----------3----------4----------5
   Very Low Medium Very High

3. **What certification level do you think you will receive?**

   Below 70% correct
   Tracker Level 1: 70-79% correct
   Tracker Level 2: 80-89% correct
   Tracker Level 3: 90-99% correct
   Senior Tracker: 100% correct
   Master Tracker: 100% plus a minimum of 15 years experience

4. **How did you learn to identify tracks and sign?** ***(Write your answers below)***

   a. Formal Training/Experience:

   b. Informal Training/Experience:

POST-WORKSHOP
November 1, 2006

<Circle Your Answers>

1. **What is your current level of confidence for correctly identifying river otter tracks?**

   1----------2----------3----------4----------5
   Very Low Medium Very High

2. **What is your current level of confidence for identifying other animal tracks and sign?**

   1----------2----------3----------4----------5
   Very Low Medium Very High

3. **What certification level did you receive?**

   Below 70% correct
   Tracker Level 1: 70-79% correct
   Tracker Level 2: 80-89% correct
   Tracker Level 3: 90-99% correct
   Senior Tracker: 100% correct
   Master Tracker: 100% plus a minimum of 15 years experience

4. **How well did you perform on the certification?**

   1----------2----------3----------4----------5
   Worse As Better
   Than Expected Expected Than Expected

5. **How well did this process measure your tracking skills?**

   1----------2----------3----------4----------5
   Very Poorly Satisfactory Very Well

6. **How satisfied are you with this workshop?**

   1----------2----------3----------4----------5
   Very Neutral Very
   Unsatisfied Satisfied

7. **Would you recommend offering this workshop to other TPWD districts?**

   1----------2----------3----------4----------5
   Definitely Not Not Sure Definitely Yes

8. **How frequently do you think tracker certifications should be offered?**

   Never Again Every year Every 2 years Every 3 yrs Other:_____

9. **Suggestions for improving the certification process?**(*Write your answer below)*

PRE-WORKSHOP
January 23 and 25, 2006

The purpose of these questions is not to evaluate you; rather it is for me to evaluate the effectiveness of the certification process. This is a completely anonymous exercise, and your name will not be used in any way.
Thanks very much for your time!
-Ciel Wharton, Texas A&M, Wildlife and Fisheries Department

<*Circle Your Answers*>

1. **What is your current level of confidence for correctly identifying river otter tracks?**

   1----------2----------3----------4----------5
   Very Low Medium Very High

2. **What is your level of confidence for identifying other animal tracks and sign?**

   1----------2----------3----------4----------5
   Very Low Medium Very High

3. **What certification level do you think you will receive?**

   Below 70% correct
   Tracker Level 1: 70-79% correct
   Tracker Level 2: 80-89% correct
   Tracker Level 3: 90-99% correct
   Senior Tracker: 100% correct
   Master Tracker: 100% plus a minimum of 15 years experience

4. **How did you learn to identify tracks and sign?** <*Circle as many as apply*>

School classes (undergrad, grad)	Outdoor/Field Experience
Field guides/Literature	Personal observations/Self-teaching
Clubs (Wildlife Soc., Boy Scouts)	Relatives (grandfather, father…)
Experience in wildlife profession	On-the-job training
Training from another biologist	TPWD otter surveys

5. **How many times have you practiced track ID since the last workshop (three months ago)?**

   0-----1-----2-----3-----4-----**5**-----6-----7-----8-----9-----**10** Other:_____

POST-WORSHOP
January 24 and 26, 2006

<*Circle Your Answers*>

1. **What is your current level of confidence for correctly identifying river otter tracks?**

   1----------2----------3----------4----------5
   Very Low Medium Very High

2. **What is your current level of confidence for identifying other animal tracks and sign?**

   1----------2----------3----------4----------5
   Very Low Medium Very High

3. **What certification level did you receive?**

   Below 70% correct
   Tracker Level 1: 70-79% correct
   Tracker Level 2: 80-89% correct
   Tracker Level 3: 90-99% correct
   Senior Tracker: 100% correct
   Master Tracker: 100% plus a minimum of 15 years experience

4. **How well did you perform on the certification?**

   1----------2----------3----------4----------5
   Worse As Better
   Than Expected Expected Than Expected

5. **How well did this process measure your tracking skills?**

   1----------2----------3----------4----------5
   Very Poorly Satisfactory Very Well

6. **How satisfied are you with this workshop?**

   1----------2----------3----------4----------5
   Very Neutral Very
   Unsatisfied Satisfied

7. **Would you recommend offering this workshop to other TPWD districts?**

   1----------2----------3----------4----------5
   Definitely Not Not Sure Definitely Yes

8. **How frequently do you think tracker certifications should be offered?**

   Never again Every year Every 2 yrs Every 3 yrs Other:_____

9. **Suggestions for improving the certification process?** <*Write answer below*>

APPENDIX C

QUESTIONNAIRE RESPONSES

Workshop I Responses

		Very Low		Medium		Very High
Q1. What is your current level of confidence for correctly identifying river otter tracks?	Pre	4	6	6	2	1
	Post	0	1	13	4	1
Q2. What is your current level of confidence for identifying other animal tracks and sign?	Pre	0	3	13	1	2
	Post	0	3	12	3	0

		Below 70%	Level 1 70-79%	Level 2 80-89%	Level 3 90-99%	Specialist 100%
Q3. What certification level do you think you will receive?	Pre	2	8	7	2	0
Q3. What certification level did you receive?	Post	13	6	0	0	0

Q4. How did you learn to identify tracks and sign? (write out your answers)

Pre

- School classes (undergrad, grad)
- Field guides/Literature
- Clubs (Wildlife Soc, Boy Scouts)
- Experience in wildlife profession
- Training from another biologist
- Outdoor/Field Exp.
- Self-teaching
- Relatives
- On-the-job Training
- TPWD otter surveys

		Worse Than Exp.		As Expected		Better Than Exp.
Q4. How well did you perform on the certification?	Post	4	6	4	4	1

		Very Poorly		Satisfactory		Very Well
Q5. How well did this process measure your tracking skills?	Post	0	0	5	5	9

		Very Unsatisfied		Neutral		Very Satisfied
Q6. How satisfied are you with this workshop?	Post	0	0	1	5	13

		Definitely Not		Not Sure		Definitely Yes
Q7. Would you recommend offering this workshop to other TPWD districts?	Post	0	0	1	4	14

		Never Again	Every Year	Every 2 Years	Every 3 Years	Other
Q8. How frequently do you think tracker certifications should be offered?	Post	0	5	10	3	1

Q9. Suggestions for improving the certification process? (write answers)	Post	Smaller Groups (11) Prior Instruction/Handout (4) Some classroom time first Less testing, more instruction Extend workshop to 1 week

Workshop II Responses

		Very Low		Medium		Very High
Q1. What is your current level of confidence for correctly identifying river otter tracks?	Pre	0	4	9	6	0
	Post	0	0	4	14	1
Q2. What is your current level of confidence for identifying other animal tracks and sign?	Pre	0	5	11	3	0
	Post	0	1	8	10	0

		Below 70%	Level 1 70-79%	Level 2 80-89%	Level 3 90-99%	Specialist 100%
Q3. What certification level do you think you will receive?	Pre	1	18	0	0	0
Q3. What certification level did you receive?	Post	2	6	10	1	0

Q4. How did you learn to identify tracks and sign? (circle as many as apply)

Pre

School classes (3)
Field guides/Literature (13)
Clubs (1)
Exp. in wildlife profession (13)
Training from another biologist (12)
Outdoor/Field Exp. (12)
Self-teaching (14)
Relatives (7)
On-the-job Training (13)
Otter surveys (10)

		Worse Than Exp.		As Expected		Better Than Exp.
Q4. How well did you perform on the certification?	Post	0	2	4	6	7

		0	1	2	3	4	5	6	7	8	9	10
Q5. How many times have you practiced track ID since the last workshop (three months ago)?	Pre	0	2	3	8	2	1	0	2	1	0	0

		Very Poorly		Satisfactory		Very Well
Q5. How well did this process measure your tracking skills?	Post	0	0	2	7	10

		Very Unsatisfied		Neutral		Very Satisfied
Q6. How satisfied are you with this workshop?	Post	0	0	0	2	17

		Definitely Not		Not Sure		Definitely Yes
Q7. Would you recommend offering this workshop to other TPWD districts?	Post	0	0	0	1	18

		Never Again	Every Year	Every 2 Years	Every 3 Years	Other
Q8. How frequently do you think tracker certifications should be offered?	Post	0	10	7	2	0

Q9. Suggestions for improving the certification process?	Post	Keep the groups small like today (2) More training periodically Review process before going to the field Great training None, excellent workshop Great just the way it is, great instructors Great job

VITA

CIEL AMY WHARTON
25 Spring Creek Rd.
Boerne, Texas 78006

EDUCATION

Master of Science, Wildlife Science, Texas A&M University, 2006

Bachelor of Arts, Environmental Studies, Prescott College, 2002

WORK EXPERIENCE

Research Technician (River Otter Scat Surveys), Missouri Department of Conservation, Houston, Missouri (May 2005 – August 2005).

Naturalist, Naturalists At Large, San Francisco Bay Area, California (September 2003 – March 2004).

Naturalist, YMCA- Great Hollow Wilderness School, New Fairfield, Connecticut (June 2003 – August 2003**).**

Research Technician (Golden-cheeked Warbler and Black-capped Vireo Surveys), Performance Group Incorporated, San Antonio, Texas (March 2003 – June 2003).

Backcountry Trip Leader, Prescott College, Prescott, Arizona (August 2002 – September 2002).

Naturalist, Great Smoky Mountains Institute at Tremont, Townsend, Tennessee (May 2002 – August 2002).

Nature Summer Camp Director and Lead Instructor, Cibolo Nature Center, Boerne, Texas (June 2001 – July 2001).