Interpretations of Polar Bear (*Ursus maritimus*) Tracks by Inuit Hunters: Inter-rater Reliability and Inferences Concerning Accuracy

P. B. Y. Wong[1,2,3], P. Van Coeverden de Groot[1], C. Fekken[4], H. Smith[5], M. Pagès,[6,7] and P. T. Boag[1]

[1] Department of Biology, Queen's University, 99 University Avenue, Kingston, Ontario K7L 3N6 Canada
[2] Department of Ecology and Evolutionary Biology, University of Toronto, 25 Willcocks Street, Toronto, Ontario M5S 3B2 Canada
[3] Department of Natural History, Royal Ontario Museum, 100 Queens Park, Toronto, Ontario M5S 2C6 Canada
[4] Department of Psychology, Queen's University, 99 University Avenue, Kingston, Ontario K7L 3N6 Canada
[5] Faculty of Education, Queen's University, 511 Union Street, Kingston, Ontario K7M 5R7 Canada
[6] Centre de Biologie pour la Gestion des Populations (Unité mixte de recherche 022 Institut national de la recherche agronomique (INRA)-Institut de recherche pour le développement (IRD)-Centre de cooperation international en recherche agronomique pour le développement (CIRAD)-Montpellier SupAgro), Campus International de Baillarguet, CS 30016, Montferrier-sur-Lezcedex 34988 France
[7] Laboratoire de génétique des microorganismes, Université de Liège, Institut de Botanique, 4000 Liège, Belgium

P. B. Y. Wong, P. Van Coeverden de Groot, C. Fekken, H. Smith, M. Pagès, and P. T. Boag. 2011. Interpretations of Polar Bear (*Ursus maritmus*) tracks by Inuit hunters: inter-rater reliability and inferences concerning accuracy. Canadian Field-Naturalist 125(2): 140–153.

Due to their tracking experience in pursuing Polar Bears (*Ursus maritimus*), Inuit hunters could provide non-invasive estimates of Polar Bear characteristics from tracks, and Polar Bear monitoring programs could benefit from Inuit input. We determined i) inter-rater reliability of estimates of the sex, age, and size of Polar Bears, and estimates of the age of tracks made by a group of nine Inuit hunters who interpreted 78 tracks; ii) we made preliminary comparisons of sex and size estimates with conventional (scientific) estimates; iii) we catalogued the Polar Bear hunting experience and track interpretation techniques of nine Inuit hunters; and iv) we explored relationships between hunting experience and the ability to interpret tracks. The group of Inuit hunters made reliable and consistent estimates of Polar Bear sex, age, and size, as well as estimates of age of track (after data from one participant was excluded). Although our comparisons are based on small samples, our findings suggest that Inuit hunters may be accurate in estimating the sex of Polar Bears (74.42% agreement with genetic determinations) and the size of Polar Bears from their tracks. Our data indicate shared tracking techniques used by hunters may explain high agreement in making specific estimates, while individual hunting experience and particular methods used to interpret tracks may lead to inter-rater reliability and accuracy in interpreting tracks.

Key Words: Polar Bear, *Ursus maritimus*, tracking, traditional knowledge, interviews, population characteristics, non-invasive population monitoring, Nunavut.

Uncertainty in range-wide responses by the Polar Bear (*Ursus maritimus*) to changes in sea ice conditions (Aars et al. 2006*; Freeman and Wenzel 2006; Dyck et al. 2007; Stirling et al. 2008) can affect contemporary estimates of sizes and dynamics of Polar Bear populations (Taylor et al. 2005; Aars et al. 2006*; Taylor et al. 2006; Dowsley 2007). Canadian Polar Bears are monitored in accordance with the 1973 Agreement on the Conservation of Polar Bears (Aars et al. 2006*; Freeman and Wenzel 2006; Stirling and Parkinson 2006) using "sound conservation practices based on the best available scientific data" (Canada et al. 1973*).

Polar Bear dynamics are mainly estimated through analyses of population viability using data from aerial capture-mark-recapture (CMR) surveys (Taylor et al. 2005; Taylor et al. 2006), which are expensive (Dowsley 2009), infrequent (Government of Nunavut 2005*; Taylor et al. 2006), and sometimes do not have the full support of local communities (Tyrell 2006; Clark et al. 2008; Shannon and Freeman 2009). The lack of accurate information on Polar Bear population dynamics can have regional consequences, ranging from erratic harvest quotas for resident Inuit (Taylor et al. 2008; Dowsley 2009) to incorrect projection of responses by Polar Bear populations to climate change and the aggravation of differences of opinion between scientific and Inuit communities (Clark et al. 2008; Dowsley 2009). A first step to a more affordable and wide-spread monitoring program may be more frequent, less invasive estimates of Polar Bear activity involving local Inuit.

The integration of local knowledge/expertise into valid scientific Polar Bear monitoring methods is a laudable goal (Agreement between the Inuit of the Nunavut settlement area and Her Majesty the Queen in right of Canada 1993*; Government of the Northwest Territories 1993*; Usher 2000), but it has proven to be elusive. Information that Inuit hunters can provide, for example, numbers, age, or sex classes of Polar Bears or location of sightings (Stirling and Parkinson 2006), has not been integrated into a rigorous repeatable

method, fuelling international criticisms of changes in quotas based on data provided by Inuit hunters (these data are considered non-scientific information) (Aars et al. 2006*). With their expert knowledge of Polar Bear morphology and behaviour (Freeman and Wenzel 2006), Inuit hunters may be the first to notice major changes in Polar Bear population dynamics (Derocher et al. 2004). They regularly predict when and where Polar Bears are likely to be at certain times of the year, and they may be able to provide valid estimates of the sex, age, and size of specific Polar Bears from their tracks. As a first step toward the inclusion of estimates from tracks into any Polar Bear activity survey, such data should be evaluated for inter-rater reliability (defined here as consistency or repeatability of estimates among individual hunters) and accuracy (the extent to which estimates are close to scientifically derived values).

Polar Bear tracks exhibit cues that are used by observers, Inuit or otherwise, to generate estimates of the sex, age, and size of the animal that made them, according to some decision rules. However, there are many sources of error in estimates made by observers, for example, lack of ability, experience, and attention, as well as fatigue. One strategy for estimating error is to assess the degree of agreement in estimates generated by multiple judges, for example, using inter-rater reliability statistics such as Ebel's intraclass correlation coefficient (Ebel 1951; Bartko and Carpenter 1976; Shrout and Fleiss 1979). Although mathematically similar to the more widely known Cronbach's coefficient alpha (Cronbach 1951; Cortina 1993; Peterson 1994; Santos 1999; Gliem and Gliem 2003*), Ebel's intraclass correlation coefficient provides an index of agreement among multiple judges (in this case, Inuit observers), while Cronbach's alpha provides an index of agreement among items (different sets of tracks). In this regard, the intraclass correlation coefficient has been used to evaluate the reliability of multiple raters in reviewing quality assessments (i.e., Kilic and Cakan 2007) and health care-related assessments (i.e., Schneider et al. 2004; Riffenburgh and Johnstone 2009; Ramachandran et al. 2011).

Beyond confirming that judges make decisions that are in agreement is demonstrating that judgments are valid or accurate relative to an external criterion. Accuracy in identifying the sex of Polar Bears from tracks can be evaluated by comparing an identification made from tracks with genetic sex determination using associated non-invasively collected tissue. Accuracy in estimating of the age of Polar Bears from tracks relies on genetically identifying animals caught in a previous aerial capture-mark-recapture study and determining their ages from tooth wear patterns (Calvert and Ramsay 1998). Inferences concerning accuracy in estimating the size of Polar Bears may be drawn by comparing the size estimate with measurements of the length of the stride from the same tracks (Heglund 1974). However, it is likely that stride length will provide more repeatable estimates of animal size in any monitoring program.

Our understanding of Inuit information about Polar Bears derived from tracks will be enhanced by considering the history and personal experience of hunters. Identifying which characteristics distinguish the most reliable and accurate hunters may also help provide subsequent recruitment criteria for Inuit hunter participation in track-based Polar Bear surveys. Preliminary assessments of the reliability of three active Inuit hunters and three elders (identified by other participants and community members) in 2007 and three active hunters and four non-Inuit in 2008 indicate that active, experienced hunters are generally more reliable observers of tracks (unpublished data).

Building on this work, we report i) inter-rater reliability of a larger group of Inuit hunters in estimating Polar Bear characteristics from tracks; ii) preliminary comparisons of Inuit hunter estimates of the sex and size of Polar Bears with external track estimates; iii) the techniques that Inuit hunters use to interpret tracks; and iv) criteria such as hunting background and tracking experience that contribute to higher inter-rater reliability and accuracy of estimates.

Methods

We randomly assigned a number (1 to 9) and distributed instructions to nine Inuit from the communities of Gjoa Haven, Taloyoak, and Cambridge Bay in Nunavut, Canada. Participants 1 to 7 identified themselves as active hunters (Inuit who presently hunt Polar Bears when provided with the opportunity), and participants 8 and 9 were elders (older community members who no longer actively hunt). Participant 1 translated all of the verbal responses of participants 8 and 9; for this reason, we collected all data from participants 8 and 9 after we had collected the data from participant 1. Participants 1 and 3 had participated in at least one of the earlier field seasons (Wong 2010*). Additional hunters were contacted through the hunters and trappers organizations in their respective communities.

Participant 1, followed by the other participants, located Polar Bear tracks via snowmobile in M'Clintock Channel between and around Cape Sydney (69°N, 97°W) and Gateshead Island (70°N, 100°W) from 11 to 15 May 2009. Sampling occurred on straight-line transects between Cape Sydney and Gateshead Island and randomly chosen transects around Gateshead Island to avoid re-sampling the tracks of individual Polar Bears. For each track, each participant provided an estimate of the sex, age in years, and nose-to-tail size in feet of the Polar Bear that had made the track, along with the age of the track in days. Participants provided their estimates without discussion with the other participants in the same order at all tracks observed. In order to minimize random guessing, we

included "I don't know" as a possible response (Yu 2001*).

Inter-rater assessment of estimates made by hunters of sex, age, and size of Polar Bears and estimates of age of track

We used group agreement, individual differences, and changes in group variability over time to determine the inter-rater reliability of the interpretations of tracks across tracks interpreted by all participants. Following MacLennan (1993), we calculated the intraclass correlation coefficient (Ebel 1951) for the group of hunters and the adjusted item-total Pearson product-moment correlation coefficients (r) for each participant (Gliem and Gliem 2003*) using SPSS (SPSS Inc.) for estimates of Polar Bear sex, age, and size, and estimates of age of track across tracks. We also determined the mean adjusted item-total r for the group by z_r-transforming the adjusted item-total r values for each participant, computing the arithmetic mean, and back-transforming mean z_r values (Silver and Dunlap 1987).

We tested for consistent differences among hunter estimates with a chi-square (χ^2) analysis of sex ratios; a one-way analysis of variance (ANOVA) for each of the age and size of Polar Bears and the age of track variables; and a post-hoc Tukey-Kramer Honestly Significant Difference (HSD) test for the age and size of Polar Bears and age of track variables (JMP version 9.0) (JMP 2009). To determine whether group variability decreased over time, we numbered tracks in the order they were interpreted and we tested the relation between track and coefficients of variation in the estimates of age and size of Polar Bears and estimates of the age of track using linear regression.

Comparisons of sex estimates with genetic determinations and comparisons of size estimates with stride measurements

To provide inferences concerning the accuracy of the estimates of sex and size of Polar Bears made by the participants, we made comparisons with respective "scientific" estimates of sex and stride length. Track data were collected alongside an ongoing study from 2007 to 2009 to genetically sex Polar Bears using non-invasive sampling stations, which consisted of a square barbed-wire fence surrounding a pole baited with seal meat (Van Coeverden de Groot et al. 2010*). These stations were erected between Cape Sydney and Gateshead Island no more than three days prior to track data collection along the same transect used to locate tracks. Samples of hair from Polar Bears that were attracted to these stations were collected between 11 and 15 May 2009 for genetic sexing and genotyping (Van Coeverden de Groot et al. 2010*) at the same time as we documented associated track interpretations made by participants (above). We also collected hair samples on an *ad hoc* basis along tracks where interpretations were made. We supplemented these data with paired track interpretations and hair samples collected in this area from 1 to 14 May 2008 and around Cape Sydney from 2 to 18 May 2007 (Figure 1).

Only hair samples that were associated with a single track (thus a single individual) were included in subsequent analyses; these associations were either obviously apparent or confirmed by participant 1. The resulting data set was composed of hair samples from 2009 associated with estimates by participants 1 to 9; samples from 2008 associated with estimates by participant 3 and one previous hunter participant from Taloyoak (participant A); and samples from 2007 associated with estimates by participants 1, 2, and A and two previous participating hunters from Gjoa Haven (participants B and C). We stored all hair samples frozen in cryovials.

Genetic sex determination of hair samples relied on a polymerase chain reaction (PCR) duplex amplification of the ZFX/ZFY and SRY markers (Pagès et al. 2009), which we first evaluated on 10 tissue samples of known sex (six females and four males). Because a faint SRY-like band was observed for some females, the initial method was modified. PCR amplicons were digested overnight at 37°C by the restriction enzyme BccI (BioLabs), which cuts ZFY into two fragments (106 bp and 37 bp) but not ZFX. This procedure was designed using the software NEBcutter V2.0 (BioLabs). Digestions were carried out in 20 µL reaction volumes containing 5 units of BccI enzyme (BioLabs) and 0.2 µL of BSA (supplied with the enzyme, 10 mg/mL). Digested amplicons were then analyzed using a microchip electrophoresis system (MultiNA, Shimadzu; and Agilent 2100 Bioanalyzer, Agilent).

A three-band profile corresponded to a male (144 bp ZFX, 115 bp SRY, 106 bp ZFY) (the 37 bp ZFY was too short to be visualized), and a one-band profile corresponded to a female (144 bp ZFX), even though a two-band profile could be observed (144 bp ZFX, faint 115 bp SRY band, but no ZFY band). The refined procedure allowed us to circumvent the amplification of false SRY in females (probably due to SOX gene family amplification), regularly observed in mammalian sexing procedures (Taberlet et al. 1993; Kohn et al. 1995; Durnin et al. 2007).

The new method was tested on 22 DNA extracts obtained from non-invasive samples (15 hair and 7 feces) of individuals of known sex (16 males and 6 females). Four of these samples failed to amplify. Reliable and accurate sex determinations were obtained for the remaining 18 samples. The sex was then determined for 33 study samples. Between one and three independent PCR attempts were carried out and analyzed to determine sex by consensus among the different attempts. In cases where multiple hair samples were associated with a single set of tracks, genetic sex determination for the sampled Polar Bear was made by consensus. We calculated the percentage of agreement for pairs of estimates of the sex of Polar Bears

made from tracks and genetic determinations from 2007 to 2009 for each hunter and across all hunters, and then, using a binomial test, we evaluated whether percentage agreement was significantly different from random guessing (50%).

In the absence of data relating stride length to body length in Polar Bears, we compared hunter estimates of size directly with stride measurements. Some participants mentioned during interviews that stride length is a useful indicator of body size, and we assumed stride length was correlated with true body size (Heglund 1974). By random assignment, in 2009 participants 4, 5, and 6 walked along each track and selected by group consensus consecutive footprints left by walking Polar Bears on flat terrain considered to be suitable for measurement. After providing their other estimates, these participants recorded a minimum of six measurements of the distance from the front edge of the left-hind footprint in one group of tracks to the front edge of the corresponding footprint in the following group at each suitable track. We compared the means of stride distances with corresponding estimates of Polar Bear size by calculating Pearson's product-moment correlation coefficients (r) for all 2009 participants.

Semi-structured interviews for hunting and track interpretation techniques and experience

We obtained permission to conduct interviews from the Nunavut Research Institute, Queen's University, and the Gjoa Haven Hunters and Trappers Organization, along with oral and written consent from each participant. Between 1 April and 11 May 2009, we conducted semi-structured interviews according to participant availability and convenience. The interviews, which occurred alone with the participant (except when a translator was required), ranged from 5 to 12 min depending on mutual comprehension and the amount of information.

All interviews were conducted following Huntington (1998, 2000). Directive questions (i.e., name, age, community) preceded general questions concerning tracking methods, with follow-up questions to allow hunters to explain their own understanding and thoughts or to clarify information (Huntington 1998, 2000; Rapley 2001). We audiotaped and transcribed all interviews, with the exception of the interview with participant 5, whose responses were mainly recorded by hand due to a malfunctioning battery in the recorder.

To provide further contextual information, we documented non-verbal cues in a journal and verbal styles through audio-recording, and we quoted each participant's experiences and reflections in his own words (Baxter and Eyles 1997; Huntington 2000; Rapley 2001). These details included data such as the participants' apparent increased comfort in answering interview questions when they were not being audiotaped. We then grouped all transcribed interviews and relevant data pertaining to track interpretation methods based on our interpretations (Burnard 1991) and used quotations and other information that best represented these groupings.

Participant background and hunting experience and inter-rater reliability and accuracy in interpreting tracks

To make a preliminary description of the relationship between the background of the participants and their ability to interpret tracks, we calculated pair-wise Spearman's rank correlation coefficients (ρ) between adjusted item-total r (for each of the sex, age, and size of Polar Bears and the age of track), the percentage agreement of sex estimates with genetic sex determinations, and r for size estimates with stride measurements and each participant's background and tracking experience.

We organized background and hunting experience using five criteria: continuous data on age and education and categorized data on frequency of guiding, preferences for hunting alone, and the ability to interpret tracks from observing few footprints. We divided frequency of guiding Polar Bear hunts into three categories ("never guides," "sometimes guides," and "often guides"); preferences for hunting alone versus hunting with a group into two categories; and the methods used to interpret tracks from observing footprints into three categories ("observing a whole track," "sometimes observing a whole track, sometimes a single footprint," and "observing only a single footprint"). We did not have the above participant background information for participant 5.

Results

The group of hunters encountered 99 tracks in total in 2009 (Figure 1). All nine participants observed 43 tracks between Cape Sydney and Gateshead Island, 25 tracks north of Gateshead Island, and 17 tracks south of Gateshead Island; participants 1 and 9 also observed 14 tracks north of Gateshead Island. All participants estimated the sex, age, and size of the Polar Bear that had made the track and the age of the track for the same set of 78 tracks. None reported "I don't know."

Inter-rater reliability of estimates of sex, age, and size, and age of track

Preliminary data analysis based on box plots revealed outlying data points that skewed the distribution of age of track estimates. Since outliers were associated with participant 5, we report inter-rater reliability assessments with and without data from participant 5 to examine his effect on inter-rater reliability. Based on interclass correlation coefficients (ICC), the group of nine participants was reliable in estimating sex, age, and size but was not reliable in estimating age of track (Table 1). Specifically, the ICC for sex, age, and size

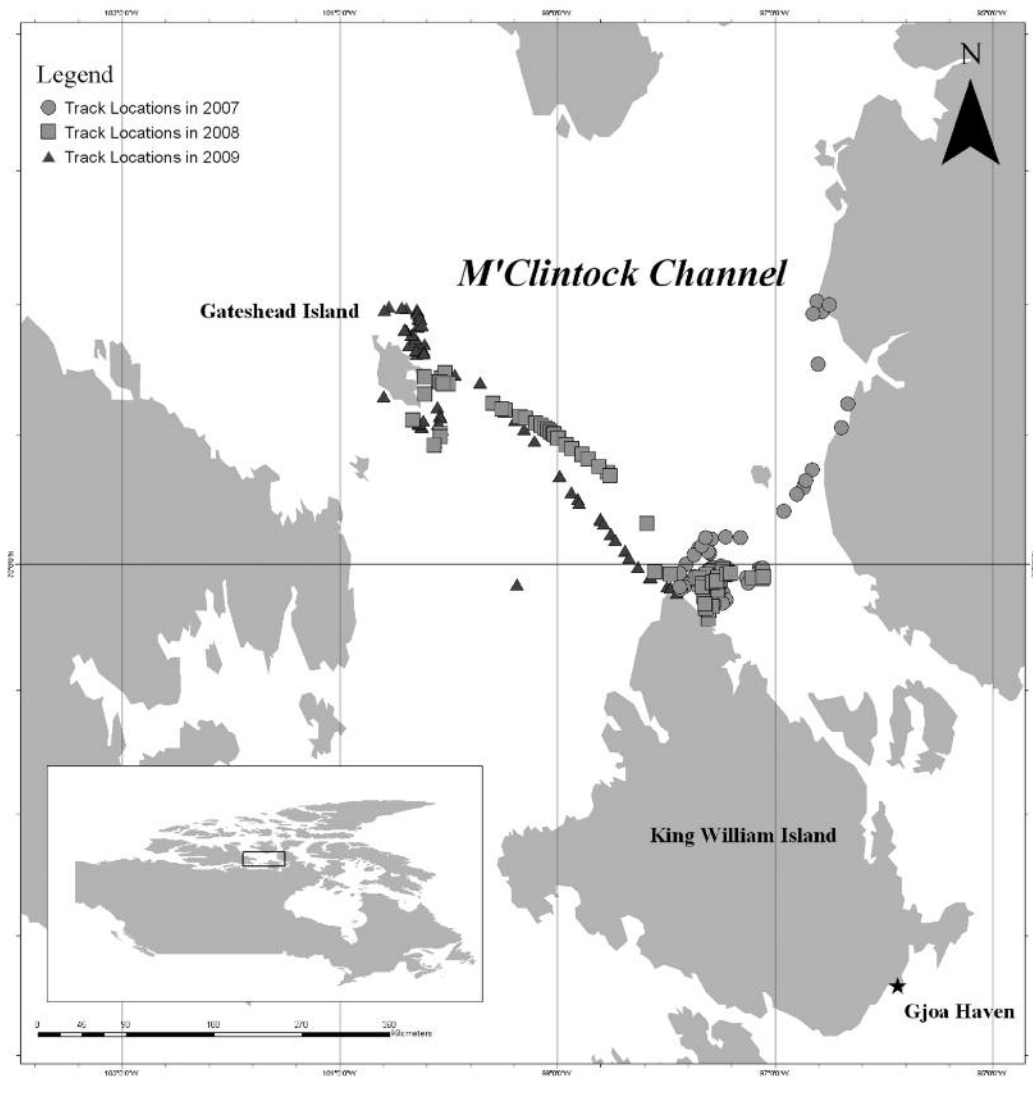


FIGURE 1. Track locations over ice in M'Clintock Channel, Nunavut, Canada, from 2007 to 2009. A total of 19 tracks were interpreted by three hunters and three elders in 2007, and a total of 27 tracks were interpreted by three hunters and four non-Inuit in 2008. In 2009 (the current study), a total of 99 tracks were encountered between Cape Sydney and Gateshead Island between 11 and 15 May. Of these 99 tracks, 78 were interpreted by all nine participants, who estimated the sex, age, and size of the Polar Bears that made them and the age of the track.

estimates exceeded 0.7, often used as the benchmark for a measurement to be considered reliable according to similar statistics (Clark and Watson 1995; Santos 1999; Streiner 2006).

When the estimates made by participant 5 are excluded, the age of track estimates are also reliable and the inter-rater reliability of estimates of sex, age, and size do not change appreciably. Based on adjusted item-total correlations between estimates made by each participant and group estimates, participants were generally consistent in making estimates of all four variables (mean adjusted item-total $r > 0.40$; Gliem and Gliem 2003*). When the estimates made by participant 5 are excluded, there is little change in consistency in the sex, age, and size estimates, although correlations of individual age of track estimates with group age of track estimates are generally higher.

Sex ratio estimates differed significantly among the nine participants ($\chi^2 = 20.20$, df = 8, $P = 0.0096$). Mean age estimates were also significantly different (one-way ANOVA, df = 8, $r^2 = 0.40$, $F = 58.95$, $P << 0.05$), with a post-hoc Tukey-Kramer HSD test

TABLE 1. Intraclass correlation coefficients (ICC) (Ebel 1951), mean adjusted item-total correlations (*r*), and individual participant adjusted item-total *r* for estimates of sex, age, and size of Polar Bears and estimates of age of track for the complete group of 9 participants and for 8 participants (excluding participant 5), made in M'Clintock Channel, Nunavut, from 11 to 15 May 2009.

Adjusted item-total Pearson product-moment correlation coefficient (*r*)

	Polar Bear							
	Sex		Age		Size		Age of track	
	All 9 participants	Participants 1–4 and 6–9	All 9 participants	Participants 1–4 and 6–9	All 9 participants	Participants 1–4 and 6–9	All 9 participants	Participants 1–4 and 6–9
ICC	0.74	0.77	0.74 ()	0.77	0.91	0.92	0.35	0.85
1	0.37	0.35	0.58	0.54	0.69	0.70	0.47	0.63
2	0.60	0.60	0.55	0.54	0.84	0.75	0.47	0.68
3	0.62	0.65	0.63	0.64	0.74	0.76	0.36	0.62
4	0.41	0.44	0.50	0.54	0.69	0.71	0.40	0.70
5	0.11		0.53		0.52		0.49	
6	0.62	0.61	0.57	0.56	0.63	0.64	0.55	0.67
7	0.33	0.35	0.81	0.79	0.82	0.82	0.64	0.76
8	0.16	0.19	0.76	0.73	0.73	0.69	0.33	0.27
9	0.64	0.62	0.63	0.62	0.78	0.76	0.44	0.55
Mean	0.45	0.49	0.63	0.63	0.73	0.74	0.47	0.63

indicating participants 3, 4, and 8 differed the most ($P << 0.05$). Mean size estimates (one-way ANOVA, df = 8, $r^2 = 0.15$, $F = 15.60$, $P << 0.05$) and age of track estimates with participant 5 (one-way ANOVA, df = 8, $r^2 = 0.07$, $F = 6.45$, $P << 0.05$) and without participant 5 (one-way ANOVA, df = 7, $r^2 = 0.07$, $F = 6.37$, $P << 0.05$) also differed significantly. Post-hoc Tukey-Kramer HSD tests indicated that participants 4, 5, and 8 differed the most in the mean estimates of size ($P < 0.05$); when participant 5 was excluded, participant 9 differed the most in the mean age of track estimates ($P < 0.05$). Relations between coefficients of variation in age of Polar Bears ($r^2 = 1.19 \times 10^{-5}$, df = 76, $P = 0.98$), size ($r^2 = 0.0025$, df = 76, $P = 0.66$), and age of track estimates, with participant 5 ($r^2 = 0.037$, df = 76, $P = 0.09$) and without participant 5 ($r^2 = 0.0037$, df = 76, $P = 0.59$), and the sequence of track observations over the study were not significant.

Comparisons of sex estimates with external determinations and comparisons of size estimates with stride measurements

We collected a total of 23 hair samples identified with sex along ten tracks in 2007, a total of 8 samples along two tracks in 2008, and a total of 22 samples along seven tracks in 2009 (Table 2); multiple hair samples were collected along 11 of these tracks. Of the total 53 hair samples, 25 samples (associated with five tracks in 2007, one track in 2008, and five tracks in 2009) were sexed in the lab and used for subsequent comparisons with hunter estimates of sex from tracks. Two samples were ambiguously sexed and 6 samples were not sexed due to repeated PCR failure, which we ascribe to too little DNA in the hair samples. Many of the samples that we failed to sex were part of a collection of samples associated with single tracks. This meant that a genetic sex determination could still be made for these tracks using other hair samples.

All hair samples were collected from Polar Bear sampling stations except for three samples that were collected along one isolated track in 2007. No hair samples were collected along multiple sets of tracks (i.e., tracks easily identifiable as females with associated cub were excluded). All participants from 2007 (participants 1 to 9), 2008 (participants 3 and A), and 2009 (participants 1, 2, A, B, and C) provided sex estimates with a mean agreement of 74.42% with the associated genetic sex determination (Table 2). These results were significantly different from a random guess frequency of 50% ($n = 41$, $P = 0.0010$).

Participants 4, 5, and 6 recorded sets of six stride measurements along each of nine tracks interpreted by all participants (2009). The time and effort required to make consistent measurements (i.e., following tracks to locate areas where the animal was walking and the left hind pad of a footprint was distinct) limited our sample, as we preferentially allocated our sampling effort to achieving a larger data set of track interpretations. In light of this limitation ($n = 9$ tracks, Table 3), estimates of animal size and measurements of stride length made by participants 1, 2 and 7 were significantly correlated. All other participant correlations were not statistically significant; however, they overwhelmingly represented a large effect size (Cohen 1992).

Semi-structured interviews to gather participant background, hunting experience, and techniques

Participants varied in terms of background and experience (Table 4): in particular, participants 1 and

TABLE 2. Percentage agreement between individual participant's estimates of the sex of Polar Bears from tracks and associated genetic sex determinations from hair collected along tracks from 2007 to 2009. The number of unique tracks with associated tissue samples used for the percentage agreement calculation for each participant is shown. The number of participants who provided sex estimates varied from 2007 to 2009. In some cases, there was more than one tissue sample associated with a track; in those cases, the consensus determination of the sex for all associated samples was used (see text for details).

Samples from 2009 were made in M'Clintock Channel between and around Cape Sydney and Gateshead Island, Nunavut, from 11 to 15 May 2009 by all nine participants. Samples from 2008 were associated with estimates by participant 3 and participant A (from Taloyoak) in M'Clintock Channel between and around Cape Sydney and Gateshead Island, 1 to 14 May 2008. Samples from 2007 were associated with estimates by participants 1, 2, and A and participants B and C (two hunters from Gjoa Haven) around Cape Sydney, 2 to 18 May 2007.

Participant	No. of track–tissue pairs	Percentage agreement
1 [a, b]	7	71.43
2 [a, b, c]	2	100.00
3 [a, b, c]	7	71.43
4 [a]	3	66.67
5 [a]	3	100.00
6 [a]	3	66.67
7 [a, b]	3	66.67
8 [d]	3	66.67
9 [d]	5	40.00
A [a]	3	100.00
B [a, b, c]	2	100.00
C [d]	2	100.00
Weighted mean		74.42

[a] Participants who actively hunt.
[b] Participants who guide Polar Bear hunts professionally.
[c] Participants who identified themselves as Canadian Rangers.
[d] Participants identified by other participants as elders.

3 were well-known professional hunters and guides; participants 2 and 3 were certified Canadian Rangers (Canada 2010*) who were also recognized as professional hunters; and participant 7 had guided Polar Bear denning surveys as well as Polar Bear hunts. On the other hand, criteria used to interpret tracks were held in common among the participants (Table 5). Sex estimates were generally made by observing footprint orientation, size, and shape. Participants indicated that male footprints were more oriented inwards, toward the centre of the track, than female footprints. Participant 2 indicated after his interview that adult male footprints are "turned in" because their shoulder muscles are more developed and the footprints of young males and females are less "turned in." Participants also indicated that footprints made by males are larger than those of females (participant 2, 5 May 2009):

TABLE 3. Pearson correlation coefficients (r) of estimates of the size (length) of Polar Bears and mean measurements of six Polar Bear stride lengths across nine tracks for all nine participants, made in M'Clintock Channel, Nunavut, from 11 to 15 May 2009.

Participant	r	P
1	0.67	0.05*
2	0.78	0.01*
3	0.54	0.13
4	0.63	0.07
5	0.65	0.06
6	0.31	0.41
7	0.69	0.04*
8	0.01	0.98
9	0.56	0.12

* Significant correlation at an alpha level of 0.05.

"And the female track and male track is different. And he said it's a different—those toes are narrow—and males are wider." (Participant 1 translating for participant 9, 12 May 2009)

Participant 1 observed accompanying tracks to identify sex as well as age:

"…Female only by herself is…5, 6 years old. Something like that…and a young male too. Same thing." (Participant 1, 10 May 2009)

Age estimates were generally dependent on footprint size or shape or on the estimated sex of the Polar Bear. All participants indicated footprint or track size differed with the age of the animal, where older Polar Bears make larger tracks. Participant 3 estimated age by inferring the weight of the animal from the depth of the track in the snow:

"…By the tracks…you can kind of tell from how big the track is… maybe the weight of the bear…depends on the snow I guess. How deep it is…." (Participant 3, 10 May 2009)

Some participants mentioned footprint shape as a cue to estimating age, with similarities between the tracks of young males and females:

"…grown up bears…tracks or footprints are more round." (Participant 2, 5 May 2009)

"…Young males'… prints are a little bit—almost same size as a female but they're more narrow." (Participant 2, 5 May 2009)

All participants examined the size of footprints to estimate the size of Polar Bears, and all participants observed weather and snow conditions or hardness and softness of prints to estimate the age of a track. Participants generally associated soft footprints with fresh tracks and hard footprints with old tracks:

"On a nice day, you can tell it's not too long ago. Even when there's drifting snow you can test it with your

TABLE 4. Summary of participants' responses to interview questions, 1 April to 11 May 2009.

Participant	Age *	Highest education level *	Place of birth	Frequency guiding hunts *	Prefers to hunt alone	Learned how to hunt from	Reasons for hunting
1	63	Grade 5	mainland	often	yes	hunters	food fur money
2	50	Grade 6	community	occasional	yes	elders father brother grandfather	food
3	51	Grade 5	community	often	yes	elders father grandfather uncle	food enjoyment
4	32	Grade 7	city	infrequent	yes	elders	food
5	no data	no data	no data	never	no data	hunters	no data
6	23	Grade 9	city	never	no	participant 7	food enjoyment
7	35	Grade 9	community	often	yes	elders hunters father	food fur money enjoyment
8	71	never attended school	mainland	infrequent	no	hunters father	food
9	67	never attended school	mainland	infrequent	no	hunters stepfather	food

* Variables compared with criteria for biases in estimates and ability to interpret tracks.

TABLE 5. Summary of criteria used by various hunters to estimate the sex, age, and size of Polar Bears from tracks and to estimate the age of tracks, based on interviews conducted between 1 April and 11 May 2009.

	Participant								
Variable	1	2	3	4	5[a]	6	7	8	9
Polar Bear									
Sex	OF, AT[b]	OF, FS	no data	OF, FS, FSh	no data	OF, FSh	OF, FSh	OF, FS, FSh	FS, FSh
Age	FS, ES, AT[b]	FS, FSh, ES	FSh, SD[b]	FS, ES	no data	FS	FS	FS	FS
Size	FS	FS	FS	FS	FS	FS	FS	FS	FS
Age of track	W, Sn	W, Sn	W, Sn	W	W, Sn	Sn	W, Sn	W, Sn	W, Sn
Method [c]	S	S	S, M	M	no data	M	M	S	S, M

OF = orientation of footprints
FS = footprint size
FSh = footprint shape
AT = accompanying tracks (i.e., male or cub tracks associated with a female)
ES = estimated sex
SD = snow depth
W = weather conditions
Sn = snow conditions
[a] Data based on notes recorded in P. Wong's journal during the interview and her recollection.
[b] Cues not mentioned by other participants.
[c] Whether hunters interpreted one or two footprints (S) or looked at a number of footprints (a track, M). These data were subsequently included in comparisons of tracking experience and inter-rater reliability and accuracy.

TABLE 6. Pair-wise Spearman's rank correlation coefficients (ρ) of participant characteristics and inter-rater reliability and accuracy of their estimates.

	Participant				
	Frequency of guiding	Age [a]	Education [a]	Frequency of accompanying hunters preferred [a]	Ability to interpret tracks from few footprints [a]
Adjusted item-total *r*					
Polar Bear					
Sex	0.06	−0.12	−0.11	0.17	−0.19
Age	0.45	0.48	−0.25	0.17	0.06
Size	0.50	0.31	−0.17	−0.28	0.32
Age of track	0.15	−0.60	0.74[b]	−0.17	−0.38
Agreement of sex estimates with genetic determinations	0.22	−0.14	0.22	−0.58	−0.51
Correlations of size estimates with stride length	0.50	−0.24	0.35	−0.73[b]	0.06

[a] Data not available for participant 5.
[b] Significance at the alpha level of 0.05.

hand, see if the footprints are hard or soft. Harder means longer and…soft means just a few hours ago or something." (Participant 2, 5 May 2009)

Comparisons of participant background and hunting experience with ability to interpret tracks

Exploratory analyses relating the characteristics of participants to indices of their inter-rater reliability and accuracy in interpreting tracks (Table 6) must be interpreted with caution; reported background and experience may not necessarily lead to reliable or accurate estimates and vice versa. Individual differences in inter-rater reliability and comparisons with external determinations were statistically defined; for example, inter-rater reliability in estimating sex was defined as the adjusted item-total correlation between a participant and all other participants for sex estimates. Moreover, the sample size for these analyses was small (nine participants), so it may be more appropriate to consider correlations that demonstrate at least a moderate effect size of .50 or greater (Cohen 1992).

More reported experience as a hunting guide was positively correlated with reliable estimates of size and experience as a hunting guide was positively correlated with more size estimates that were correlated with external determinations. Older participants were more reliable in estimating the age of Polar Bears but were less reliable in estimating the age of tracks. More educated participants showed greater inter-rater reliability in estimating the age of tracks. Preferring to hunt with other hunters was associated with estimating sex and size in lower agreement and lower correlation with external determinations, respectively. Claiming to be able to gain information from relatively few tracks was associated with lower agreement of sex estimates with genetic determinations.

Discussion

Inter-rater reliability of Inuit hunters in estimating Polar Bear characteristics from tracks

This research provides promising findings regarding the potential value of information provided by Inuit hunters, particularly their estimates of Polar Bear features given only Polar Bear tracks to observe. A group of nine Inuit hunters generally agreed in their estimates of sex, age, and sex of Polar Bears. Moreover, over time, there was no change in group variability in estimating these characteristics. For a small number of tracks, Inuit hunters' sex estimates were in higher agreement with genetic determinations than with random guessing. Overall, these findings, along with variation among hunter estimates, were enriched by relating hunters' self-reported background and hunting experience to inter-rater reliability and the correlations of their estimates with scientific determinations. As might be anticipated, more hunting experience and confidence were associated with greater degrees of inter-rater reliability and comparability with externally determined identifications.

We expected that a group of hunters with shared knowledge and experience in tracking would show high agreement. However, our group was reliable in estimating the age of tracks only after the data from one participant (5) were excluded. Low inter-rater reliability can result from random guessing (Yu 2001*); to increase inter-rater reliability, participants must be re-examined for their experience or removed from subsequent surveys (Peterson 1994; Santos 1999). Alternatively, low inter-rater reliability in estimating the age of tracks suggests the effect of track degradation and substrate should be investigated. Although collective information about tracks through group discussion can be highly accurate (Stander et al. 1997),

TABLE 7. Comparisons of track interpretation techniques used by Inuit hunters and published scientific data on Polar Bears.

Inuit track interpretation technique	Scientifically reported categorical difference
Larger track sizes of males versus females • orientation of footprints toward centre of track in males due to larger "shoulder" muscles; • similarity between female tracks and young male tracks.	Extended growth period in males results in greater body length and mass in adult males than in females of similar age (Derocher and Wiig 2002; Derocher et al. 2005); average footprint width of 22 cm in adult males and 16 cm in adult females (Amstrup et al. 2006).
Using snow or weather conditions to determine age of track.	Track degradation or age of track increases with recent snow fall (Amstrup et al. 2006), wind, ice overflow, and melt-out (Hayward et al. 2002).
Using the presence of accompanying tracks to estimate sex and age of Polar Bear.	Accompanying tracks indicate maturity, births, and disappearance of young (Jewell et al. 2001).
Using inferences of body weight from depth of tracks in the snow to estimate the age of the Polar Bear.	Body mass and age are correlated in Polar Bears (Derocher et al. 2005).

our study shows high inter-rater reliability of estimates made by a group of hunters, suggesting hunters may be able to reliably interpret tracks alone.

Preliminary comparisons of estimates of Polar Bear sex and size from tracks with external determinations

Our data suggest that Inuit made estimates in agreement with genetic determinations 74% of the time. Although this value may be criticized due to the small sample size of track interpretations with associated tissue, it is the only instance of comparing Inuit interpretations of tracks with external validity criteria. We recognize that tracks may vary in probability of accurate identification, for example, easily identifiable tracks belonging to females and their associated cubs and large tracks likely belonging to adult males are easier to interpret. Although easily identifiable tracks were excluded from our analyses, a larger sample of track and tissue pairs (excluding easily identifiable tracks) will be required to provide inferences concerning the accuracy of Inuit estimates of the sex of adult Polar Bears.

Genetic sexing of Polar Bear hair is still a relatively new procedure (Pagès et al. 2009), and optimizing techniques for feces and blood will provide genetic determinations from additional pairs of tracks and tissue for validation (Van Coeverden de Groot et al. 2010*). Work to genotype non-invasive tissue associated with previously captured (Taylor et al. 2006) and aged (Calvert and Ramsay 1998) Polar Bears to compare with estimates of age made by hunters from tracks is also warranted, given the high inter-rater reliability of the Inuit estimates reported herein. Calibrations of body size and stride length in zoo animals or "rogue" Polar Bears detained in Churchill (Tyrell 2006) will be required to establish a robust relationship between animal size and stride length in Polar Bears for further comparison with estimates made from tracks. At a minimum, all sampling efforts (both wild and known Polar Bears) should include measurements taken along tracks made in similar snow substrate that minimizes variability in gait patterns (i.e., walking versus galloping). Distinguishing among Polar Bears of similar age and sex (identity) awaits additional data such as the optimization of multivariate analyses of digital images of Polar Bear tracks (Alibhai and Jewell, personal communication; Jewell et al. 2001; Alibhai et al. 2008), and will be required for a valid estimate of Polar Bear abundance (Hayward et al. 2002; Silveira et al. 2003).

At the outset, a number of limitations were associated with a survey of animal tracks based exclusively on interpretations of tracks made by hunters. Gathering track data requires high densities of individual animals (Smallwood and Fitzhugh 1995) and minimal variance in individual travel distances (Stephens et al. 2006). Animals must be mobile in order to be counted (Becker et al. 1998); otherwise, a failure to detect tracks will not necessarily indicate absence (Gese 2001; Crooks et al. 2008). In addition to physical geography (Gompper et al. 2006; Houser et al. 2009), weather and snow conditions can affect track detectability (Jewell et al. 2001; Hayward et al. 2002; Silveira et al. 2003; Alibhai et al. 2008), thus further supporting the need to evaluate the effect of substrate on interpretations made by Inuit hunters.

Against these concerns, our data indicate Inuit have the potential to provide information on more than simply presence or absence of specific Polar Bears in an area. Coupled with genetic data from remote sampling stations (Harris 2010*), this information could provide preliminary estimates of contemporary Polar Bear activity that are currently unavailable for most populations. Before this information is included, however, evaluations of accuracy in estimating Polar Bear age and size and the age of tracks need to be completed. It is important to note that there are no data on individual and inter-rater reliability and accuracy in hunter estimates of sex, age, and size for any target animal species (but see Stander et al. 1997). Once

these evaluations have been completed, the challenge will be to integrate Inuit track data into a valid Polar Bear activity survey with genetic data from non-invasive samples (Harris 2010*) and potentially digital images from tracks.

Polar Bear population characteristics estimated from tracks alone could provide repeatable objective data (Smallwood and Fitzhugh 1995; Stander et al. 1997; Gusset and Burgener 2005). However, the margins of error in ongoing capture-mark-recapture surveys continue to be smaller. This means that capture-mark-recapture surveys result in more accurate harvest quotas that would presumably minimize over- or under-harvesting. At this stage, the track data discussed here should supplement rather than replace capture-mark-recapture data, especially when capture-mark-recapture surveys or direct visualization of animals is difficult (Beier and Cunningham 1996; Jewell et al. 2001).

Tracks alone can be used to detect the presence or absence of individual Polar Bears (Balme et al. 2009); behavioural activity such as births, disappearance of young, maturity, and maternity (Jewell et al. 2001); and, more importantly, large changes in population activity over large areas (Balme et al. 2009) and time (Kendall et al. 1992; Hayward et al. 2002; Melville and Bothma 2006; Balme et al. 2009) with high statistical power in detecting decreases in abundance (Beier and Cunningham 1996). These data could, in conjunction with modelling techniques associated with capture-mark-recapture surveys, inform managers when another survey is required to re-evaluate harvest quotas. It is important to note that data derived from tracks will require independent estimates of population parameters, such as those provided through genetic analyses or capture-mark-recapture, to be applied to management (Herzog et al. 2007).

Hunting and tracking experience as indicators of inter-rater reliability and potential accuracy in interpreting tracks

Hunters indicated that they acquired their tracking skills through interactions with a similar group of people (elders, other hunters, and family members). Although the small sample size limited statistical comparisons between information from interviews and estimates made by participants, our qualitative interpretations suggest inter-rater reliability of hunters in interpreting tracks may be explained by their shared track interpretation techniques. Many of the shared skills reported by Inuit can be directly linked to Polar Bear characteristics identified in conventional scientific studies (Table 7). We anticipate that optimizing multivariate analyses of digital track images will provide morphometric data against which these shared criteria (Table 5) can be directly compared. On the other hand, some hunters (participants 1 and 3) identified unique track interpretation techniques, perhaps acquired as a function of their extensive tracking experience or their involvement in other scientific research projects (in 2007 and 2008). These few, context-rich cases can identify particular track interpretation techniques that might be taught to a larger group (Maxwell 2004; Flyvbjerg 2006) should their techniques be arguably better for interpreting tracks.

The qualitative data also explore individual differences in Inuit estimates of Polar Bear characteristics from tracks. Participant 3 is an experienced hunter, and the wider range in his age estimates may accurately reflect the true range in Polar Bear age, which has been estimated to be up to 20 years in female Polar Bears (Ramsay and Stirling 1988). Likewise, participants 4 and 5 showed the greatest differences in estimates from the rest of the hunters (in addition to participant 5 lowering the effect on inter-rater reliability); this difference may be ascribed to their less frequent participation in Polar Bear hunts. Participant 8 (an elder) also showed differences in his estimates as well as lower agreement with other hunters and external determinations; this difference may be explained by his general lack of active hunting. Participants 1, 2, and 7 (all experienced professional guides) gave estimates of animal size most correlated with stride length. The hunters whose size estimates correlated the least with stride length were again the elder (participant 8) and a young hunter (participant 6). Overall, our study suggests Inuit from any community who show levels of expertise comparable to the most reliable and likely accurate hunters reported here (see above) could provide useful non-invasive estimates of Polar Bear characteristics from tracks for any population.

Acknowledgements

Funding was provided by the Natural Sciences and Engineering Research Council of Canada, the Nunavut Wildlife Management Board, Nunavut Tunngavik Inc., Sigma Xi Grants-in-aid-of-Research, and the Northern Scientific Training Program. Molecular work was supported by a fellowship to M. Pagès from the FRS-FNRS (Fonds de la Recherche Scientifique – Fonds national de la Recherche Scientifique) (chargée de recherches) in Belgium. This project was possible only with the Gjoa Haven Hunters and Trappers Organization and its members and members from the Taloyoak Hunters and Trappers Organization and the Cambridge Bay Hunters and Trappers Organization who provided data, knowledge, and safety for this study. C. Harris and T. Vording assisted in data collection. H. Smith, S. Lougheed, S. Arnott, and C. Eckert provided guidance and assistance in data interpretation and analysis.

Documents Cited (marked * in text)

Aars, J., N. J. Lunn, and **A. E. Derocher.** 2006. Proceedings of the 14th Working Meeting of the IUCN/SSC Polar Bear Specialist Group. Seattle, Washington, 20-24 June 2005. Occasional Paper of the IUCN Species Survival Commission. International Union for Conservation of Nature and Natural Resources, Gland, Switzerland.

Agreement between the Inuit of the Nunavut settlement area and Her Majesty the Queen in right of Canada. 1993.

Canada. 2010. Canadian rangers. Department of National Defence. http://www.army.forces.gc.ca/land-terre/cr-rc/index-eng.asp (accessed November 2010).

Canada, Denmark, Norway, the Union of Soviet Socialist Republics, and the United States of America. 1973. Agreement on the Conservation of Polar Bears. 15 November 1973. Oslo, Norway.

Gliem, J. A., and **R. R. Gliem.** 2003. Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales. 2003 Midwest Research to Practice Conference in Adult, Continuing, and Community Education 82-88.

Government of Nunavut. 2005. Polar bear management memorandum of understanding between Cambridge Bay, Ekaluktutiak Hunters' and Trappers' Organization; Gjoa Haven, Gjoa Haven Hunters' and Trappers' Organization; Taloyoak, Spence Bay Hunters' and Trappers' Organization; Kitikmeot Hunters' and Trappers' Association and the Department of Environment for the management of the "M'Clintock channel" polar bear population. Nunavut Department of Environment.

Government of the Northwest Territories. 1993. Traditional Knowledge Policy. Government of the Northwest Territories, Yellowknife.

Harris, C. 2010. Noninvasive capture-recapture techniques to manage and conserve polar bears. Master's thesis, Department of Biology, Queen's University, Kingston, Ontario.

Van Coeverden de Groot, P. J., C. Harris, P. Wong, M. Dyck, and **P. Boag**. 2010. Report to the Gjoa Haven Hunters and Trappers Organization on Polar Bear fieldwork 2006–2009. Queen's University, Kingston, Ontario.

Wong, P. 2010. Reliability, accuracy, and tracking techniques of Inuit hunters in estimating polar bear characteristics from tracks. Master's thesis, Department of Biology, Queen's University, Kingston, Ontario.

Yu, C. H. 2001. An introduction to computing and interpreting Cronbach coefficient alpha in SAS. Pages 1-7 *in* Proceedings of 26th SAS User Group International Conference, Long Beach, California, April 22-25, 2001.

Literature Cited

Alibhai, S. K., Z. C. Jewell, and **P. R. Law**. 2008. A footprint technique to identify white rhino *Ceratotherium simum* at individual and species levels. Endangered Species Research 4: 205–218.

Amstrup, S. C., I. Stirling, T. S. Smith, C. Perham, and **G. W. Thiemann.** 2006. Recent observations of intraspecific predation and cannibalism among polar bears in the southern Beaufort Sea. Polar Biology 29: 997–1002.

Balme G.A., L. T. B. Hunter and **R. Slotow.** 2009. Evaluating methods for counting cryptic carnivores. Journal of Wildlife Management 73: 433-441.

Bartko, J. J., and **W. T. Carpenter.** 1976. On the methods and theory of reliability. Journal of Nervous and Mental Disease 163: 307–317.

Baxter, J., and **J. Eyles.** 1997. Evaluating qualitative research in social geography: establishing 'rigour' in interview analysis. Transactions of the Institute of British Geographers 22: 505–525.

Becker, E. F., M. A. Spindler, and **T. O. Osborne.** 1998. A population estimator based on network sampling of tracks in the snow. Journal of Wildlife Management 62: 968–977.

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

Burnard, P. 1991. A method of analysing interview transcripts in qualitative research. Nurse Education Today 11: 461–466.

Calvert, W., and **M. A. Ramsay.** 1998. Evaluation of age determination of polar bears by count of cementum growth layer groups. Ursus 10: 449–453.

Clark, D. A., D. S. Lee, M. M. R. Freeman, and **S. G. Clark.** 2008. Polar bear conservation in Canada: defining the policy problems. Arctic 61: 347–360.

Clark, L. A., and **D. Watson.** 1995. Constructing validity: basic issues in objective scale development. Psychological Assessment 7: 309–319.

Cohen, J. 1992. A power primer. Psychological Bulletin 112: 155–159.

Cortina, J. M. 1993. What is coefficient alpha? An examination of theory and applications. Journal of Applied Psychology 78: 98–104.

Cronbach, L. J. 1951. Coefficient alpha and the internal structure of tests. Psychometrika 16: 297–334.

Crooks, K. R., M. Grigione, A. Scoville, and **G. Scoville.** 2008. Exploratory use of track and camera surveys of mammalian carnivores in the Peloncillo and Chiricahua mountains of southeastern Arizona. Southwestern Naturalist 53: 510–517.

Derocher, A. E., and **Ø. Wiig.** 2002. Postnatal growth in body length and mass of polar bears (*Ursus maritimus*) at Svalbard. Journal of Zoology 256: 343–349.

Derocher, A. E., N. J. Lunn, and **I. Stirling.** 2004. Polar bears in a warming climate. Integrative and Comparative Biology 44: 163–176.

Derocher, A. E., M. Andersen, and **Ø. Wiig**. 2005. Sexual dimorphism of polar bears. Journal of Mammalogy 86: 895–901.

Dowsley, M. 2007. Inuit perspectives on polar bears (*Ursus maritimus*) and climate change in Baffin Bay, Nunavut, Canada. Research and Practice in Social Sciences 2: 53–74.

Dowsley, M. 2009. Community clusters in wildlife and environmental management: using TEK and community involvement to improve co-management in an era of rapid environmental change. Polar Research 28: 43–59.

Dowsley, M., and **G. Wenzel.** 2008. "The time of the most polar bears": a co-management conflict in Nunavut. Arctic 61: 77–189.

Durnin, M. E., P. J. Palsbøll, O. A. Ryder, and **D. R. McCullough.** 2007. A reliable genetic technique for sex determination of giant panda (*Ailuropoda melanoleuca*) from non-invasively collected hair samples. Conservation Genetics 8: 715–720.

Dyck, M. G., W. Soon, R. K. Baydack, D. R. Legates, S. Baliunas, T. F. Ball, and **L. O. Hancock.** 2007. Polar bears of western Hudson Bay and climate change: are warming spring air temperatures the "ultimate" survival control factor? Ecological Complexity 4: 73–84.

Ebel, R. 1951. Estimation of the reliability of ratings. Psychometrika 16: 407–424.

Flyvbjerg, B. 2006. Five misunderstandings about case-study research. Qualitative Inquiry 12: 219–245.

Freeman, M. M. R., and **G. W. Wenzel.** 2006. The nature and significance of polar bear conservation hunting in the Canadian Arctic. Arctic 59: 21–30.

Gese, E. M. 2001. Monitoring of terrestrial carnivore populations. Pages 372–396 *in* Carnivore Conservation, *edited by* J. L.Gittleman, S. M. Funk, D. W. Macdonald, and R. K. Wayne. Cambridge University Press, Cambridge.

Gompper, M. E., R. W. Kays, J. C. Ray, S. D. Ladpoint, D. A. Bogan, and **J. R. Cryan.** 2006. A comparison of noninvasive techniques to survey carnivore communities in northeastern North America. Wildlife Society Bulletin 34: 1142–1151.

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

Hayward, G. D., D. G. Miquelle, E. N. Smirnov, and **C. Nations.** 2002. Monitoring Amur tiger populations: characteristics of track surveys in snow. Wildlife Society Bulletin 30: 1150–1159.

Heglund, N. C. 1974. Scaling stride frequency and gait to animal size: mice to horses. Science 186: 1112–1113.

Herzog C. J., R. W. Kays, J. C. Ray, M. E. Gompper, W. J. Zielinski, R. Higgins and **M. Tymeson.** 2007. Using patterns in track-plate footprints to identify individual fishers. Journal of Wildlife Management 71: 955-963.

Houser, A. M., M. J. Somers, and **L. K. Boast.** 2009. Spoor density as a measure of true density of a known population of free-ranging wild cheetah in Botswana. Journal of Zoology 278: 108–115.

Huntington, H. P. 1998. Observations on the utility of the semi-directive interview for documenting traditional ecological knowledge. Arctic 51: 237–242.

Huntington, H. P. 2000. Using traditional ecological knowledge in science: methods and applications. Ecological Applications 10: 1270–1274.

Jewell, Z. C., S. K. Alibhai, and **P. R. Law.** 2001. Censusing and monitoring black rhino (*Diceros bicornis*) using an objective spoor (footprint) identification technique. Journal of Zoology 254: 1–16.

JMP. 2009. Version 9. SAS Institute Inc., Cary, N.C.

Kendall, K. C., L. H. Metzgar, D. A. Patterson, and **B. M. Steele.** 1992. Power of sign surveys to monitor population trends. Ecological Applications 2: 422–430.

Kilic, G. B., and **M. Cakan.** 2007. Peer assessment of elementary science teaching skills. Journal of Science Teacher Education 18: 91–107.

Kohn, M., F. Knauer, A. Stoffella, W. Schröder, and **S. Pääbo.** 1995. Conservation genetics of the European brown bear—a study using excremental PCR of nuclear and mitochondrial sequences. Molecular Ecology 4: 95–104.

MacLennan, R. N. 1993. Interrater reliability with SPSS for Windows 5.0. American Statistician 47: 292–296.

Maxwell, J. A. 2004. Causal explanation, qualitative research, and scientific inquiry in education. Educational Researcher 33: 3–11.

Melville H. I. A. S., and **J. D. Bothma.** 2006. Using spoor counts to analyse the effect of small stock farming in Namibia on caracal density in the neighbouring Kgalagadi Transfrontier park. Journal of Arid Environments 64: 436-447.

Pagès, M., C. Maudet, E. Bellemain, P. Taberlet, S. Hughes, and **C. Hänni.** 2009. A system for sex determination from degraded DNA: a useful tool for palaeogenetics and conservation genetics of ursids. Conservation Genetics 10: 897–907.

Peterson, R. A. 1994. A meta-analysis of Cronbach's coefficient alpha. Journal of Consumer Research 21: 381–391.

Ramachandran, V., J. D. Lewis, M. Mosstaghimi-Tehrani, B. A. Stach, and **K. L. Yaremchuk.** 2011. Communication outcomes in audiologic reporting. Journal of the American Academy of Audiology 22: 231–241.

Ramsay, M. A., and **I. Stirling.** 1988. Reproductive biology and ecology of female polar bears (*Ursus maritimus*). Journal of Zoology 214: 601–633.

Rapley, T. J. 2001. The art(fulness) of open-ended interviewing: some considerations on analysing interviews. Qualitative Research 1: 303–323.

Riffenburgh, R. H., and **P. A. Johnstone.** 2009. Measuring agreement about ranked decision choices for a single subject. International Journal of Biostatistics 5: Article 17.

Santos, J. R. A. 1999. Cronbach's alpha: a tool for assessing the reliability of scales. Journal of Extension 37. http://www.joe.org/joe/1999april/tt3.html (accessed July 2008).

Schneider, B., K. Maurer, D. Sargk, H. Heiskel, B. Weber, L. Frölich, K. Georgi, J. Fritze, and **A. Seidler.** 2004. Concordance of DSM-IV Axis I and II diagnoses by personal and informant's interview. Psychiatry Research 127: 121–136.

Shannon, K. A., and **M. M. R. Freeman.** 2009. Inuit observations of polar bears in Salliq/Coral Harbour, Nunavut, and the management of the conservation hunt. Pages 39–50 *in* Inuit, Polar Bears, and Sustainable Use, *edited by* M. M. R. Freeman and L. Foote. CCI Press, Edmonton, Alberta.

Shrout, P. E., and **J. L. Fleiss.** 1979. Intraclass correlations: uses in assessing rater reliability. Psychological Bulletin 86: 420–428.

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.

Silver, N. C., and **W. P. Dunlap.** 1987. Averaging correlation coefficients: should Fisher's *z* transformation be used? Journal of Applied Psychology 72: 146–148.

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.

SPSS. 2009. Version 17.0.1. SPSS Inc., Chicago, Illinois.

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

Stephens, P. A., O. Y. Zaumyslova, D. G. Miquelle, A. I. Myslenkov, and **G. D. Hayward.** 2006. Estimating population density from indirect sign: track counts and the Formozov-Malyshev-Pereleshin formula. Animal Conservation 9: 339–348.

Stirling, I., and **C. L. Parkinson**. 2006. Possible effects of climate warming on selected populations of polar bears (*Ursus maritimus*) in the Canadian Arctic. Arctic 59: 261–273.

Stirling, I., A. E. Derocher, W. A. Gough, and **K. Rode**. 2008. Response to Dyck et al. (2007) on polar bears and climate change in western Hudson Bay. Ecological Complexity 9: 193–201.

Streiner, D. L. 2006. Starting at the beginning: an introduction to coefficient alpha and internal consistency. Journal of Personality Assessment 80: 99–103.

Taberlet, P., H. Mattock, C. Dubois-Paganon, and **J. Bouvet.** 1993. Sexing free-ranging brown bears *Ursus arctos* using hairs found in the field. Molecular Ecology 2: 399–403.

Taylor, M. K., P. D. McLoughlin, and **F. Messier.** 2008. Sex-selective harvesting of polar bears *Ursus maritimus*. Wildlife Biology 14: 52–60.

Taylor, M. K., J. Laake, P. D. McLoughlin, E. W. Born, H. D. Cluff, S. H. Ferguson, A. Rosing-Asvid, R. Schweinsburg, and **F. Messier.** 2005. Demography and viability of a hunted population of polar bears. Arctic 58: 203–214.

Taylor, M. K., J. Laake, P. D. McLoughlin, H. D. Cluff, and **F. Messier.** 2006. Demographic parameters and harvest-explicit population viability analysis for polar bears in M'Clintock Channel, Nunavut, Canada. Journal of Wildlife Management 70: 1667–1673.

Tyrell, M. 2006. More bears, less bears: Inuit and scientific perceptions of polar bear populations on the west coast of Hudson Bay. Inuit Studies 30: 191–208.

Usher, P. J. 2000. Traditional ecological knowledge in environmental assessment and management. Arctic 53: 183–193.

Received 11 June 2011
Accepted 24 August 2011