

Supporting Information: Punishment sustains large-scale cooperation in pre-state warfare

Sarah Mathew and Robert Boyd

Contents

1	Recruitment and administration of study	2
2	Population mortality (data reported in Figure 1)	4
2.1	Data collection	4
2.2	Possible biases in mortality estimates	5
2.3	Life stage versus age	7
2.4	Classification of mortality causes	8
3	In-depth interviews (data reported in Figures 2–7)	8
3.1	Data collection	8
3.2	Categorizing raids into stealth versus force	11
3.3	Raids with over 1000 people	11
3.4	Data transformation procedures	12
3.5	Possible biases in estimates	13
3.6	Missing data	15
3.7	Inter-subject variation (data presented in Tables S1 and S2)	16
3.8	Confidence Interval Estimation	17

4	Vignette Study (data reported in Figure 8)	17
4.1	Scenario for raiding another ethnic group	20
4.2	Scenario for raiding another territorial section	21
4.3	Questionnaire and instructions	21
4.4	Vignette data analysis	23
5	Reports of large-scale combat in hunter-gatherers	24

1 Recruitment and administration of study

The data were collected during 9.5 months of field work by Sarah Mathew from 2008–2010. Participants were a representative sample of adult men reliant on nomadic pastoralism for their livelihood. We recruited them in a town close to the ethnic border frequented by nomadic Turkana who live in the surrounding 50 km radius. Recruitment was done with the help of trained local Turkana research assistants. The RA approached potential participants, briefly introduced them to the study, and then asked them to participate. If they agreed, they were brought to the study site described below. We conducted the study in the town center because the risk of raids in the surrounding nomadic settlements was high. We did not recruit from the settled Turkana population in this town center as they are Turkana who have actively pursued schooling and professions outside the pastoral sector, or were forced out of subsistence pastoralism after droughts, epidemics or raids. Nomadic Turkana frequent the perimeter of the town center for a variety of reasons as it is an important trading center and is adjacent to dry season watering holes. Herders from surrounding nomadic settlements frequent the market to sell an animal and purchase tobacco, tea,

sugar and flour. Turkana women come to sell milk and collect relief food. Nomadic Turkana periodically come to town to visit their settled relatives. A river runs along the perimeter of town. The wells dug in the dry bed of the river provide water for surrounding nomadic settlements and so herdsman frequently bring their livestock for watering to the perimeter of town from where we can recruit them. To ensure that our sample is representative of herders within the town perimeter on a given day, we aimed to recruit using the following procedure to the extent that circumstances allowed: First, we specified the age-group that the participant should belong to. Then the RA went to one of the locations where nomadic herders visiting town are found. He approached the first person of the pre-specified age-group that he saw that he could reasonably approach. A potential participant was considered as being “reasonably approachable” if the RA could approach him without causing offense, rudely interrupting an ongoing conversation, or bringing undue attention to the act. If a person refused to participate, the RA repeated the process. We cycled through multiple age-groups and recruiting locations, to prevent drawing participants from the same network cluster. Circumstances did not always allow this procedure when there were too few members of an age-group around on a particular day. In these cases, after making a reasonable attempt, the RA picked the next available nomadic herder from any age-group he saw.

Virtually all Turkana in our potential subject pool have experienced an offensive raid first-hand at least once. So we did not have to rely on snowball sampling in the study except to recruit multiple individuals who participated in the same raid (to estimate inter-subject variation; see Section 3.7).

All studies were conducted at the study site—a room located between the market area and the watering holes. This prevented onlookers and an audience of peers,

which minimized biases towards norm-compliant responses. Participants were compensated for their time with 200 Ksh (3 USD) for vignette studies lasting around 30 minutes–1.5 hours and 500 Ksh for the interviews which lasted 3–6 hours. Studies were administered in the Turkana language. A local Turkana research assistant was present along with the researcher. For the semi-structured interviews, questions were framed by the researcher, the RA then asked the participant the question in Turkana, and translated the participant’s response immediately to the researcher, before the next question was asked. A digital recording of each participant’s response was also obtained and a fraction of these were translated independently by another local research assistant. Gaps in the first translation were filled using the second translation from the digital recording when possible. For the vignette study, the RA narrated the scenario and questions to the participant in the presence of the researcher.

2 Population mortality (data reported in Figure 1)

2.1 Data collection

Population mortality was estimated by reconstructing the family structure of participants. Among the 113 male participants recruited in the interview study, 6 participants did not meet internal consistency checks while responding to the demographic questions, and so were dropped from the study. Participants listed all their offspring and full siblings. We asked for all live births in these categories to be reported. Three reported still births were excluded. For each live birth that was listed we asked whether the person was still alive. If they were, we obtained a description of their current life stage. If they had died, we asked how they died, and asked for a description of their life stage at the time of their death. Participants, and their wives and parents

were not included in the sample to avoid bias in the mortality estimates. Wives and parents are drawn from the population who survived to reproductive age. Parents are more likely to be individuals who have survived long in their reproductive life stage and produced more children, thus contributing more to the participant pool. Using only offspring and siblings of participants provides an unbiased estimate of mortality to the extent that deaths occur independently within a family. Interviews of 107 participants yielded a sample of 446 males across the three life stages (222 in the middle childhood to puberty life stage, 86 in the puberty to reproductive life stage, and 138 in the reproductive life stage). We did a similar study with 125 female participants (64 married and 61 unmarried women) . Four of these participants did not meet internal consistency checks and so were dropped from this study. We asked women about their full siblings and offspring when the schedule allowed. Otherwise, married women were asked only about their offspring, and unmarried women were asked about their full siblings. This yielded a sample of 276 males across the three life stages (113 in the middle childhood to puberty life stage, 88 in the puberty to reproductive life stage, and 75 in the reproductive life stage).

2.2 Possible biases in mortality estimates

Mortality estimates can be biased if people are unwilling to talk about family members who have died, or if they tend to forget children who have died. Both biases would cause us to under-estimate mortality. If these factors are more likely to affect children who died in disease rather than raids, it will lead us to over-estimate raid-related mortality. Here we discuss the mitigation of these factors in our estimates.

Grief for relatives who have died prevents people from bringing up these facts if they have no reason to. Therefore, when asking people about family structure, we

acknowledged to them that we understand this may be difficult, but that it is one of the issues that our study is focused on. Therefore it was important for us to know of all their live births (or the live births of their mother or wives). We found that most participants were forthcoming and willing to talk about those who died after we framed the discussion this way.

As a further precaution, we used multiple methods of counting to check for consistency in their responses. First, participants were asked to list the total number of live births. Then they listed the total number of male and female children and the total number of children who are living and who have died. Then they listed the total number of living male, living female, died male and died female children. This allowed us to tally their totals in two ways, and check between these two for consistency. Finally, they were asked to list all the children in the order of their birth, telling us if they are male or female, living or died. We then asked them to provide the life stage for each living child, and for each child who died we asked about the cause of death and life stage at which they died. While this method does not solve the problem that children who are truly forgotten will consistently not be mentioned, it provided participants multiple opportunities to recollect and it gave us the chance to resolve inconsistencies in their accounts.

Combat deaths are more noteworthy and so parents and siblings may be less likely to forget individuals who died in raids than from illnesses. It is also plausible that younger siblings come to hear of older brothers who died in battle from stories their parents and others tell them, but may not hear of ones who died of illness. However, we think under-reporting of disease mortality due to this will be strongest for infant and early childhood deaths, not adult mortality. For parents, death of adult children will on average be more recent than death of infants, and so is less likely to be

forgotten. Also, losing an older child is more uncommon than losing an infant, and so may be more memorable. Also, the longer a child lived the more likely the sibling we interviewed was around at the time the individual died. For all these reasons we think that this bias will be strongest for infant and early child mortality rather than the life stages that we report in Figure 1.

Our child mortality estimates do not indicate strong under-reporting of disease mortality. Our data show that 49% of male children and 46% of female children die from illness before they reach middle childhood life stage (i.e. before reaching about 6 years of age). This falls within the range of child mortality in natural fertility populations without reliable access to medical care, suggesting that the extent of under-reporting of disease mortality must not be very strong. If it was not very strong for infancy and early childhood, then it is even weaker for disease mortality at older ages for reasons we discussed earlier.

2.3 Life stage versus age

We used life stage to approximate age categories, because participants could reliably provide life stage information for their family members, but could not reliably report the age or the Turkana calendar year of birth of all their family members. The start of middle childhood represents the age at which children transition from the stage when a child is mostly playing in the homestead, to one in which the child begins to leave the homestead to do chores. Thus, boys were categorized as being in middle childhood if they had started to herd goat kids on their own. This transition occurs around 4–7 years of age, depending on the child’s development, household needs and number of other siblings around. When an individual diverged prominently from the life stage of others of his age, he was categorized by his age, not life stage. For

instance, if a man cannot sire children but other men of his age can, he is categorized as being in the *reproductive* life stage, rather than as a *post reproductive* male.

2.4 Classification of mortality causes

Deaths that occurred in the course of a raid, but that were not induced by the enemy, were not classified as mortality from raids. These included four cases of death due to thirst (or exhaustion) during the travel to or from the raid, and one case of a killing of one Turkana by another in the course of a quarrel that ensued over loot division (which was classified as internal violence).

3 In-depth interviews (data reported in Figures 2–7)

3.1 Data collection

We used in-depth semi-structured interviews to elicit a detailed account of events that occurred in a raid. Interviews were the only feasible means to obtain data on what happened during raids because raids cannot be safely or legally observed. Interviews were structured in three stages. First, we asked demographic questions about the participant's family composition, age-group, clan, territorial section, settlement, degree of reliance on the pastoral economy, livestock wealth, recent migration trajectory, history of participation in raids, livestock losses from raids, and frequency of participation in group defense. Second we asked questions aimed at reconstructing the participant's offensive and defensive raiding history. After obtaining a chronological ordering of the last few offensive raids that the participant had joined, we obtained an elaborate account of his most recent offensive raid. We did this because the par-

ticipant is more likely to accurately remember details of his most recent raid and also because it ensures that our data are not biased towards raids that were particularly memorable. The interview is semi-structured and we ask questions that allow us to re-construct a detailed picture of the events that occurred. Among other topics, each participant described how the raid was planned, how combatants were mobilized, how many warriors joined, the range of settlements from which people came, where and when they attacked, how the battle unfolded, whether there were desertions, whether men lagged behind others or did not fire their weapon during combat, how many people were injured and killed, how many livestock were acquired, and what happened to the men who were identified as deserters and cowards.

We obtained first-hand accounts from 118 participants which yielded data on 88 different offensive raids. This sample was built in two stages. In the first stage, using the recruitment procedure detailed in Section 1, we recruited 113 participants. Six participants did not meet internal consistency checks while responding to the demographic questions, and so were dropped from the study. Among the 107 remaining participants, we asked 11 for a detailed summary of their raiding history and this did not leave time for them to elaborate on their most recent raid. Thus 96 participants were asked to give a detailed account of their most recent raid. Of these, two had never joined an offensive raid, and one had last joined an offensive raid about 30 years ago and could not provide specifics about the event. Thus, we have 93 participants who provided us a detailed account of their most recent raid. The data file for one of these 93 participants was inadvertently overwritten. Among the 92 participants remaining, the most recent raid of 12 participants overlapped and they provided information only on four unique raids. (These 12 participants are the informants of the first four raids in Tables S1 and S2). Thus we have detailed first hand accounts of 84

different raids. Eighty three of these raids are the most recent stealth or force raid the participant joined. One raid is not. The participant was keen to discuss this raid in particular because he had become engaged in a prolonged fire-fight with a person he knew in the opposition. While personally memorable to the subject, the raid was not unique in other ways and so we have retained it.

The second stage was a follow-up study in which we used snowball sampling to recruit multiple individuals who had participated in the same force raid. Our goal was to evaluate inter-subject reliability and variation in force raid accounts. Twenty-six participants were recruited for this study and they provided first-hand accounts of four unique force raids (the last four raids in Tables S1 and S2).

When we had multiple accounts of one raid, we used the following system to aggregate the data: we averaged the values from the multiple participants for the number of fatalities, loot gained, size of the raiding party, number of settlements, territorial sections and age-groups, and number of close kin. We used the maximum value across participants to estimate the proportion of raids in which desertion or cowardice was observed and proportion of raids in which desertion or cowardice was sanctioned. Thus, for instance, if at least one participant reported that they observed cowardice, the raid was categorized as one in which cowardice was observed. If at least one participant reported that individuals identified as cowards were sanctioned, the raid was categorized as one in which cowardice was sanctioned. For three other variables—whether or not the loot division system failed, whether combat occurred or not, and whether it was a stealth or force raid, there was no variation among subjects.

3.2 Categorizing raids into stealth versus force

Raids where number of participants were fewer than 20 men were prototypical of stealth raids, and raids where number of participants were 50 or more were recognizably force raids. Ten of the 88 raids in our sample were intermediate in size comprising 20–50 participants. We categorized these as stealth or force using the following rule: if the raiding party initiated or intended to initiate an exchange of fire, it was categorized as a force raid unless the participant explicitly said that this was a stealth mission. Using this schema, 4 of the 10 intermediate sized raids were categorized as force raids—the party size of these raids were 28, 30, 47 and 49 warriors. The raiding party size of the remaining 6 raids were 22, 28, 30, 30, 40 and 46 warriors. They were categorized as stealth raids because participants were explicit that this was the strategy. In 4 of these 6 stealth raids (raiding party size of 28, 30, 40 and 46) the raiding party initiated fire-exchange.

3.3 Raids with over 1000 people

Five raids where participants estimated the size of the raiding party to be over 1000 people were excluded from the size estimates as these reported estimates were not consistent with other measures of the size of the raiding party (like the spatial span when seated or walking and number of subgroups involved). These 5 raids are also excluded from the casualty rate estimates because calculating the casualty rate required knowing the size of the raiding party. Data from the multiple accounts of a single raid (described in section 3.7) also indicate that estimates in the range of a few thousand are not in line with what other informants said, and so 5 such estimates were excluded when aggregating the data for each raid (see Table S1). Other estimates from these 10 participants were retained because these participants are not outside the range of

what other informants say with regards to their other assessments (see Table S1).

We do not think excluding these estimates causes us to underestimate the size of the raiding party by much. Rather, we think that the people who provided these estimates were not in a position to assess accurately for one of several reasons. They may have joined the raiding party in the late stages of the gathering, or even after the raiding party began walking, and therefore were absent when some form of numerical assessment was made. Often, while the raiding party is gathering, the numerical strength is assessed by counting various subgroups who have arrived and who cluster themselves by sitting under different trees. These clusters are typically age-group based, but when the geographical span is large and the numbers high, people also aggregate according to the areas and territorial sections they come from. So, for instance, the assessment is made by saying, 50 men from this settlement have come, 30 men of this age-group, 20 from another age-group, and so on, which provides a rough estimate for the total gathering. Individuals who arrive at this gathering at late stages, either because they heard the news late, or they had to travel somewhere to find a family member to take over herding duties, or made visits to friends to borrow food and bullets for the raid, would have missed this. Their estimates may thus be based on simply estimating the size of a crowd. Under these situations, we think that those individuals who reported very high numbers and whose estimates of size were inconsistent, gave qualitatively incorrect numerical estimates.

3.4 Data transformation procedures

When participants provided a range for the size of the raiding party, we took the midpoint of the highest and lowest value in the range.

Livestock gained was primarily in the form of cows. When goats and donkeys were

obtained, they were converted into cows using the exchange rate prevalent in north Turkana.

When an estimate of the minimum is given (denoted in Table S1 with a “+” following the estimate), we used the minimum as the actual estimate. For instance, suppose the informant says that 3 people were killed from his settlement, and he also heard that there were people from other settlements who died, but he does not know the number. We denote this in the table as “3+”, and in the data analysis this is replaced with “3”. This procedure biases some estimates as we discuss next.

3.5 Possible biases in estimates

The estimates for the per capita casualty rates and number of settlements in a force raid are likely to be low estimates. For the number of Turkana who died in a raid, participants may not always have known of casualties among warriors from other territorial sections or far away settlements, especially if the raiding party did not regroup for loot division. Also, casualties that occurred at the beginning of a battle are more widely known because combatants are closer together. Deaths also occur in the later stages of combat, as the raiding party has begun retreating and combatants are dispersing with the enemy in pursuit. Warriors may not hear about casualties from the later stage of the battle if they and the victim are from different territorial sections. Nonetheless, loss of life is a key focus of conversation after raids, and so more often than not, people eventually hear the reports from other territorial sections and distant settlements about how many were killed. So, we do not expect this bias to be too large.

For the number of settlements, participants were able to refer to individual settlements from their own localities using the name of the leader of that settlement. However,

when people came from more distant lands, particularly outside of the territorial section of the participant, the participants often referred to the geographic areas these men hailed from, rather than their settlements. This means, if more than one settlement comes from the same geographic region, the participant may not know if there was only one, or whether there were actually two or three different settlements there. Thus our estimates of the number of settlements represented in a raiding party may be smaller than the actual.

The estimates of desertion and cowardice rate were obtained by asking the interviewee to report on the behavior of fellow combatants. This method could lead to over-reporting of free-riding if the interviewee exaggerated the failure of others in order to enhance his own performance. However, we do not think that this resulted in a substantive over-estimation in our study. Several factors contribute to our assessment. One, if impression-management is occurring, it is unclear whether it will lead to over or under reporting. It is true that being better than his peers increases a man's status. But it is also true that a man's status is linked to the status of his age-mates. Consequently, men are proud not only of their personal performance but also that of their age-mates in battle. Age-groups are collectively shamed and praised for the cowardice and bravery of their men in combat. Putting down one's age-mates is not a characteristic means to raise one's status. These motivations may cause participants to under report rather than over report the failures of their age-mates. Two, we framed the questions relating to cowardice and desertion without highlighting the comparative aspect of their and his conduct. Participants were not asked to evaluate their own "performance" in the raid, only to narrate on the course of events as it unfolded. Thus, we tried to treat desertion and cowardice as events, rather than an evaluation of self and other. Three, participants were not just providing a Yes/No

answer to the questions. They were providing a narrative account. When they talked about cowardice they also had to describe the events that comprised it: what the violator’s act was, in what stage of combat did it occur and how others reacted to it. On occasions when participants would speak in vague generalities, they were reminded that we wanted to hear about events and persons specific to this particular raid. Four, overall participants were very forthcoming during the interview, and freely talked about their lack of success in getting any loot, in the defeat of the raiding party and of raids that were aborted half-way without combat. One participant told us that he was a coward on that raid and was being encouraged by the rest. One participant told us about having returned on the way due to an illness. Another participant told us how he was punished. This gives us some grounds to think that the elaboration of what happened on the raid was not perceived by participants as an opportunity for bragging. Finally, we think that a more significant factor is that participants could not be aware of all the desertions and cowardice that happen in large force raids involving hundreds of warriors. They are aware of what happens in a smaller group—typically men they were standing alongside, age-mates, and settlement members. This is why despite the concern of over-reporting, we think that the estimates for the proportion of raids in which desertions and free-riding occur are under-estimates, and should be interpreted as a lower bound.

3.6 Missing data

In some instances, data from a participant is missing. In many of these cases, time constraints prevented us from completing the interview in full. Various factors gave rise to such constraints—time of day, how long the previous participant took, how long earlier stages of the interview of the participant took, and whether the participant

had to leave in a hurry in order to reach his settlement before dusk. In other cases, participants were not in a position to observe or hear about the detail being asked them. For instance, if a participant arrived at the raid just as it was departing, he may not know the various settlements from which people arrived. If the participant was injured in the raid and had to be brought home early, he may not know about cowardice in combat. If the participant's homestead was far from the rest, he may have left for home soon after and not been around to hear of sanctions, or casualties. In these cases, the missing data do not bias the estimate. The one case where we expect missing data from a subject to bias the estimate is for the number of settlements. Some participants said that there were "many" settlements in that raid, but could not give a number. Such estimates were treated as missing data in our analysis. It is likely that estimates of "many" were given for raids in which people came from a large number of settlements causing us to underestimate.

3.7 Inter-subject variation (data presented in Tables S1 and S2)

Tables S1 and S2 show the estimates provided by different participants who talked about the same raid. (Section 3.1 describes how this sample was compiled). The raw estimates of size (provided in Table S1) show what the participants have said. The transformed estimate shows what entered the data analysis. When estimates of over 1000 were provided, see Section 3.3 for the procedure we followed. For desertions and cowardice and sanctions against them (Table S2), even when multiple participants in one raid reported that they observed cowardice and desertions, participants were each referring to a different set of individuals. When participants only gave an estimate that there were "many" (for instance in Table 1, under the column "# of settlements"),

this is treated as an n/a in the data analysis. Items coded as “d/k” (Table S1) indicate that the participant responded that he does not know (i.e. he is aware that he does not possess the accurate state of information). Blank cells in Tables S1 and S2 are missing data because we did not ask the question either due to time constraints, or because it was evident that the participant could not have known about this particular aspect. In the case of sanctioning, participants were not asked the sanctioning question when they were referring to cowardice or desertions for which they would not have been in a position to know whether sanctioning occurred. For instance, it may have been too early after the raid for punishment to have been meted, the violators were from another area, or the subject only vaguely heard that some other wing had some incidence and does not know much more. For the last two participants of Raid 7, we could not finish the full interview due to time constraints.

3.8 Confidence Interval Estimation

For the per-capita fatality per raid and per-capita loot gained per raid, the data were not normally distributed. So we did an ordinary non-parametric bootstrapping procedure, and calculated the confidence interval of the mean using the bias-corrected accelerated method.

4 Vignette Study (data reported in Figure 8)

Participants in the vignette study are told a short hypothetical story. Then we ask them to narrate the story back to us. This ensures that the participant understood the story, and also allows us to examine what aspects of the story the participant finds most relevant. We then ask participants a series of questions that elicit their

Table S1: Multiple accounts of the same raid. Each row shows the scale and fatality estimates of one interviewee. Our estimate of the time that lapsed between the raid and the interview is shown in the “Recency” column.

Raid	Recency	Size (raw)	Size (transformed)	# of territorial sections	# of age-groups	# of settlements	# died
1	1 day	200	200	3	4	many	d/k
	2 days	200	200	3	4	3	d/k
	3 days	130	130	1	4+	3	0
	1 month	100	100	3	5	3	1
	1 month	250	250	3	4	4+	1
	3.5 months	104	104	2	2	3	1
2	4 days	500	500	4	5	many	0
	5 days	250	250	4	3	many	0
3	12 years	1000+	1000	3	6	3	80
	12 years	900	900	4	d/k	3	38+
4	23 years	d/k	n/a	4	6	4	0
	23 years	1260+	1260	3	6	3	0
5	9 months	200	200	4	7	6	2
	9 months	d/k	n/a	5	11	6	7
	9 months	350	350	4	5	5	7
	9 months	500 or 1000	n/a	5	7	7	7
	9 months	3000	n/a	4	7	7	6
	9 months	240	240	4	6	6	20
	9 months	d/k	n/a	3	5	3+	2
	9 months	400-500	450	4	7	7	4+
	9 months	4000-5000	n/a	4	7	8	7
	9 months	300	300	6	6	9	2+
6	5 months	250	250	5	7	10	5
	5 months	280	280	3	7	8	5
	5 months	4000	n/a	3	5	9	5
	5 months	250	250	2+	7	4	5
	5 months	250	250	3	7	5	3
	5 months	200	200	4	6	8	2
7	3 days	200, 240-300	250	5	8	9	0
	3 days	150	150	4	7	10	0
	3 days	4000-5000	n/a	6	8	8	1
	3 days	300	300	d/k	8	9	1
	3 days	250	250	4	8	8	1
	3 days	200	200	6	9	9	0
	3 days	200	200	3			1
	3 days	300	300				1
8	10 months	100	100	3	5	4	2+
	10 months	140	140	d/k	3	4	3

Table S2: Multiple accounts of the same raid. Each row shows the reports of one interviewee to whether free-riding and sanctioning of free riders occurred, and the share of the livestock they obtained on the raid.

Raid	Desertion	Desertion Sanctioned	Cowardice	Cowardice sanctioned	Loot division failed	Per capita gain
1	no	n/a	yes		n/a	0
	no	n/a	no	n/a	n/a	0
	yes		yes		n/a	0
	no	n/a	yes	yes	n/a	0
	yes	no	yes		n/a	0
	no	n/a	no	n/a	n/a	0
2	no	n/a	no	n/a	n/a	0
	no	n/a	no	n/a	n/a	0
3	yes	no	yes	yes		2
	yes	no	no	no		15
4	yes	no	no	no	no	10 cows, 20 goats
	yes	no	no	no	no	48 cows
5	no	n/a	no	n/a	yes	0
	no	n/a	no	n/a	yes	0
	no	n/a	no	n/a	yes	8 cows
	yes	no	no	n/a	yes	0
	no	n/a	yes	yes	yes	1 cow
	no	n/a	no	n/a	yes	4 cows
	yes	no	yes	no	yes	1 cow
	no	n/a	no	n/a	yes	2 cows
	no	n/a	yes	yes	yes	0
	yes	yes	no	n/a		0
6	yes		yes		yes	5 cows, 3 donkeys, 8 goats
	yes	no	yes	yes	yes	5 cows
	yes	no	yes	yes	yes	8 cows, 2 donkeys
	yes	no	yes	yes	yes	7 cows
	no	n/a	no	n/a	yes	2 cows
	no	n/a	no	n/a	yes	5 cows
7	no	n/a	n/a	n/a	yes	2 cows
	yes	yes	yes	yes	yes	2 cows
	yes	yes	yes	yes	yes	2 cows
	no	n/a	yes	yes	yes	2 cows
	yes	yes	yes	yes	yes	2 cows
	no	n/a	yes	yes	yes	0
8						0
	yes	yes	yes	no	yes	5 cows
	yes	yes	yes	no	yes	5 cows

judgment of the character and reaction towards his act.

Each question has two stages—an “open response” stage, and a “forced choice” stage. We first obtain the open response, in which the participant freely elaborates in his response to the question. To ensure that we do not let participants talk until we have heard what we want to hear, we consider an answer to be complete when the participant pauses for 3–5 seconds. The 3–5 second time interval was estimated by the RA by silently counting up to three. If the participant hasn’t begun speaking in this interim, we ask him if there is anything else he wants to add. If he says that there isn’t, we proceed to the second stage. Otherwise, we continue to record his response, until another 3–5 second pause. In the forced choice stage, we ask him to decide one way or the other—for instance, to decide whether it is correct or wrong. We use the two-stage method because participants may provide fairly complex judgments in the open response stage, positing that there were elements of the act that were justified and elements that were wrong. In the second stage we force a choice, which allows us to infer which direction the participant leans towards. In cases where the participant is still ambivalent in the forced choice stage, we repeat the forced choice question once more. If he is still undecided, we proceed to the next question.

4.1 Scenario for raiding another ethnic group

Once upon a time there was a Turkana man called Etabo from Mogilla. One day, Etabo said to his friends – “why can’t we go and look for animals in Toposa land.” So they left. When they reached the land of the Toposa, they found Toposa cows grazing, being herded by two shepherds. So they fired their weapons, killed the shepherds, and raided all the animals. They drove the animals back to Mogilla where they came from. This is the story of Etabo.

4.2 Scenario for raiding another territorial section

Once upon a time there was a Turkana man called Lopeyok from Mogilla. One day, Lopeyok said to his friends – “why can’t we go and look for animals in Lukumong land.” So they left. When they reached the land of the Lukumong, they found Turkana cows grazing, being herded by two shepherds. So they fired their weapons, killed the shepherds, and raided all the animals. They drove the animals back to Mogilla where they came from. This is the story of Lopeyok.

4.3 Questionnaire and instructions

Now I want to read you a story, and afterwards hear from you about what you’ve heard in the story. There is no correct answer that we are looking for, other than your own opinion. Have you understood? Listen carefully.

Once upon a time there was a Turkana man called Etabo from Mogilla. One day, Etabo said to his friends – "why can't we go and look for animals in Toposa land." So they left. When they reached the land of the Toposa, they found Toposa cows grazing, being herded by two shepherds. So they fired their weapons, killed the shepherds, and raided all the animals. They drove the animals back to Mogilla where they come from. This is the story of Etabo.

Have you understood the story or do you want to hear it again?

You still remember the name of the character? Who?

Now, can you repeat the story that I’ve just read to you?

Now we want to hear from you about the Etabo I just read to you about. You will be asked a question to which you should answer at length, not briefly. Thereafter a

second question will follow and you will choose briefly, saying whether it is like that, or it is not like that.

1 What Etabo did was correct or wrong? Why?

Now, choose one: it is correct or wrong?

2 Are you pleased or displeased with Etabo? Why?

Now, choose one: you are pleased OR displeased

3 Will Etabo be praised or criticized? How?

Now, choose one: he will be praised OR criticized

4 Will you say anything to Etabo if you meet him? What?

Now, choose one: you will tell OR you won't tell

5 Should anything be done to Etabo? What?

Now, choose one: Something will be done OR Nothing will be done

If yes: Who should do this? How?

6 Should his age-mates advise him? How?

6 Now, choose one: He should be advised OR He should not be advised

7 Will Etabo change his behavior in the future? How?

Now, choose one: He will change OR He won't change

8 Will you tell others about what Etabo did? What?

Now, choose one: You will tell OR You won't tell

9 Will you be next to Etabo in a raid? Why?

Now, choose one: You will be beside OR You won't be beside

10 Will you trust Etabo to herd your animals? Why?

Now, choose one: You will trust OR You won't trust

11 Will you lend Etabo your animal for future repayment? Why?

Now, choose one: You will lend OR You won't lend

12 If Etabo wants to marry your daughter, will you accept? Why?

Now, choose one: You will accept OR You won't accept

In the second vignette that the participants listen to, we use the same procedure but skip the long instructions. Instead we just say, “Now I’ll read you another story. Listen carefully.” We alternated the order in which the stories were narrated, such that half the participants heard the “raid another ethnic group” vignette first, and the other half of the participants heard the “raid another territorial section” vignette first.

4.4 Vignette data analysis

The data in Figure 8 show participants’ response in the forced choice stage to 8 of the 12 questions that were asked—questions 1-3, 5, and 9-12. The remaining 4 questions are ones in which an affirmative or negative response in the forced choice stage does not signify the motivation to sanction or not. We will present the results from qualitative data analyses on these responses and the open response of the other questions in a future publication.

We had one undecided response in the forced choice (in the *raid another ethnic group* condition) and this was coded as 0.5 (where Yes = 1 and No = 0).

One participant (from the *raid another territorial section* condition) was excluded because he gave contradictory responses in the open response and forced choice stage in 8 of the 11 questions, and this was likely due to confusion about why the question was being repeated twice. A few participants (10%) reverse their opinions on at least one of the questions. We did not exclude these responses as they did this for only 1–3 questions at most, and it seemed that a change of mind rather than confusion was behind such reversals. These reversals occurred in both directions (switching from wanting sanctions to not, and vice versa).

A between-subject comparison yields similar results to the within-subject comparison shown in Figure 7.

5 Reports of large-scale combat in hunter-gatherers

[1] reports that Russian explorers to the Aleutian islands were attacked by Alutiiq war parties comprising 200 warriors armed with bows, arrows and shields in one attack and up to 340 warriors in another attack (p. 39). Prior to contact, there was ongoing warfare between the Unangan and Alutiiq, and entire settlements could be destroyed during raids between them. Based on ethno-historic and oral historical reports, Maschner summarizes (p. 39) that “large numbers of men participated in warfare with several hundreds of individuals involved in some battles.”

[2] discusses a first-hand account of a battle that occurred in early 18th century between the Shoshones and the Blackfoot Indians. This battle occurred prior to the arrival of horses and guns. The informant recounts that there were 350 warriors on

their side, and the size of the Shoshone opposition seemed larger to him (pp. 34–35).

[3] discusses a battle involving 700 participants that took place between people inhabiting the Clarence River Valley in Australia (p. 59).

[4] discusses a raid in which 86 Yurok raided the the Hupa, and the return raid involved “about 100 raiders” (pp 50–52). There is also an account of a Mojave raid against the Apache involving 200 Mojave and a number of other groups (pp. 751–53). These groups had a few horses.

[5] reports that inter-societal conflict between the Eskimo and Athapaskan speakers in Western Alaska included “open battles with dozens or (rarely) even hundreds of combatants” (p. 18). More than one village could mobilize for defense. Burch reports that in one raid “a large invading force reportedly was spotted far enough from its target to give runners time to recruit reinforcements from another village” (p. 20).

[6] reports that Cree Indians living near the Albany river would join forces with the ones living near Moose river, to launch attacks against the Inuit. He reports of such a raid that occurred in 1736, where “Seventeen canoes from the Albany River area and eight from the Moose River area embarked on this expedition, suggesting a party of approximately fifty warriors (assuming two persons per canoe)” (p. 45).

[7] discusses an account of a party of 9 Spanish soldiers who were attacked by a group of 60 Chumash warriors (p. 91). Later on, when one of the 60 men was captured, he recounted the composition of the party, which reveals that they were mobilized from six different villages (pp. 92–93). In another encounter, Johnson reports that “a party of at least 58 men from Castec and 7 other interior Chumash and Yokuts rancherias” attacked a party of 28 people.

[?] describes a prominent battle between two nomadic foraging populations in the Central Australian Desert—the Walbiri and Waringari. The battle was over a few

water wells and the territory around it that the Waringari were using. Meggitt describes that “...in a pitched battle for the possession of the water the Walbiri drove the Waringari from the area, which they incorporated into their own territory. By desert standards the engagement was spectacular, the dead on either side numbering a score or more” (p. 42).

[8] reports on an attack that wiped out a whole camp of Aranda, a nomadic hunter-gatherer population in the Central Desert in Australia. Strehlow describes (p. 125), “A large party of avengers drawn from the Matuntara area along the Palmer River, and from some Southern Aranda local groups, was accordingly assembled...”. All but three people in the camp were killed by the attackers.

[9] discusses a raid that occurred in 1875 in the Central Desert in Australia where 80–100 men, women and children were killed by a raiding party of 50–60 warriors (p. 163).

References

- [1] Maschner HDG, Reedy-Maschner KL (1998) Raid, retreat, defend (repeat): the archaeology and ethnohistory of warfare on the North Pacific Rim. *Journal of Anthropological Archaeology* 17:19–51.
- [2] Secoy F (1953) *Changing Military Patterns on the Great Plains* (University of Nebraska Press, Lincoln).
- [3] Lourandos H (1997) *Continent of Hunter-Gatherers: New Perspectives in Australian Prehistory* (Cambridge University Press).

- [4] Kroeber AL (1976) *Handbook of California Indians* (Dover Publications, New York).
- [5] Burch ES (2007) *North American Indigenous Warfare and Ritual Violence*, eds Chacon RJ, Mendoza RG (University of Arizona Press), pp 11–29.
- [6] Bishop CA, Lytwyn VP (2007) *North American Indigenous Warfare and Ritual Violence*, eds Chacon RJ, Mendoza RG (University of Arizona Press), pp 30–57.
- [7] Johnson JR (2007) *North American Indigenous Warfare and Ritual Violence*, eds Chacon RJ, Mendoza RG (University of Arizona Press), pp 74–113.
- [8] Strehlow TGH (1970) *Australian Aboriginal Anthropology: Modern Studies in the Social Anthropology of the Australian Aborigines*, ed Berndt RM (University of Western Australia Press), pp 92–140.
- [9] Kimber RG (1990) *Hunter-Gatherer Demography: Past and Present*, eds Meehan B, White N (University of Sydney Press, Sydney), pp 160–168.