


# Bonobos use call combinations to facilitate inter-party travel recruitment

Isaac Schamberg<sup>1</sup>  · Dorothy L. Cheney<sup>2</sup> · Zanna Clay<sup>3</sup> · Gottfried Hohmann<sup>4</sup> · Robert M. Seyfarth<sup>1</sup>

Received: 4 October 2016 / Revised: 13 March 2017 / Accepted: 17 March 2017  
© Springer-Verlag Berlin Heidelberg 2017

## Abstract

Many primates produce vocalizations when initiating travel. These “travel calls” are often acoustically similar to vocalizations unrelated to travel, and listeners appear to rely on a shared context with callers to correctly interpret the calls. When individuals use vocalizations to coordinate movement with out-of-sight group mates, however, such pragmatic cues are unavailable. Under these circumstances, effective communication may depend on more informative acoustic signals. Here, we investigate travel-related vocalizations that occur when callers and listeners cannot see one another: long-distance calls given by wild bonobos (*Pan paniscus*). We find that production of a specific call combination, the “low hoot-high hoot,” is more likely than a high hoot alone to be produced prior to travel. Furthermore, the low hoot-high hoot combination is more likely to result in inter-party recruitment—that is, individuals from other parties are more likely to approach the caller. We also compare these observations

with previous research and find that bonobos appear to use distinct call combinations to facilitate specific movement patterns common in fission-fusion social structures. These results suggest that use of call combinations allow bonobos to convey more specific information than do single call types alone and that this additional information allows for effective communication between out-of-sight parties.

## Significance statement

When an animal hears a conspecific vocalization, it is able to respond appropriately by integrating information about the call type itself and the context in which it was given. Vocalizations produced by out-of-sight individuals, therefore, present a challenge: how can a listener respond appropriately with only partial information about the context of the call? In situations where distant bonobos are communicating with one another, we find that call combinations are potentially more informative to listeners than single call types alone. The use of call combinations in such situations may allow listeners to respond appropriately to conspecific vocalizations even with ambiguous information about the context in which the call was produced.

Communicated by R. I. M. Dunbar

**Electronic supplementary material** The online version of this article (doi:10.1007/s00265-017-2301-9) contains supplementary material, which is available to authorized users.

✉ Isaac Schamberg  
sisaac@sas.upenn.edu

**Keywords** Bonobo · Call combination · Fission-fusion · Travel coordination · Vocal communication

## Introduction

Vocalizations often function to facilitate group travel in primate groups (review of New World monkeys: da Cunha and Byrne 2009; review of Old World monkeys: Fischer and Zinner 2011). In many species, individuals signal their motivation to move by producing vocalizations prior to traveling (Boinski and Garber 2000). As more individuals begin to call,

<sup>1</sup> Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104-6018, USA

<sup>2</sup> Department of Biology, University of Pennsylvania, Philadelphia, PA 19104-6018, USA

<sup>3</sup> Department of Psychology, University of Birmingham, Birmingham, UK

<sup>4</sup> Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

the likelihood that the group will begin to travel increases. For example, mountain gorillas (*Gorilla beringei beringei*) increase production of grunts before the initiation of travel, and these grunts may serve as a mechanism to assess collective motivation to move (Stewart and Harcourt 1994). Similarly, in chacma baboons (*Papio hamadryas ursinus*), the likelihood that a group will begin to travel increases as the number of individuals producing grunts increases (Fischer and Zinner 2011). Other mammalian species also produce calls prior to group movement. Among meerkats (*Suricata suricatta*), for example, individuals begin to travel once a threshold of approximately three calling individuals has been achieved (Bousquet et al. 2011).

In several species, the vocalizations of particular individuals, rather than the total number of individuals who are vocalizing, appear to influence group travel. In white-faced capuchins (*Cebus capucinus*), single individuals use trill vocalizations both to initiate travel and to change travel direction (Boinski 1993). Similarly, individual African elephants (*Loxodonta Africana*) attempt to initiate group travel by producing rumbles in combination intention movements (Poole 2011). Finally, chimpanzees (*Pan troglodytes*) produce “travel hoos” prior to bouts of travel; these vocalizations appear to be targeted at particular individuals and signal a caller’s motivation to travel jointly with closely bonded individuals (Gruber and Zuberbühler 2013).

The “travel” vocalizations of these species differ in many respects. One feature they share, however, is that the vocalizations associated with travel are acoustically similar to calls produced in other contexts. Baboons, for example, produce grunts both prior to travel and as signals of benign intent when approaching others. Playback experiments have demonstrated that while baboons distinguish between the two call subtypes, context also influences listeners’ responses (Rendall et al. 1999; see also Wheeler and Fischer 2012, Price et al. 2015, and Seyfarth and Cheney 2016a on the importance of context in call interpretation). Similarly, bonobos (*Pan paniscus*) produce “travel” peeps that are acoustically similar to peeps given in other contexts (Clay et al. 2015). Thus, while vocalizations appear to play an important role in initiating travel in many species, the vocalizations themselves may only be interpreted as “travel” signals when listeners are able to integrate other contextual information, including the behavior of nearby animals.

If “travel” vocalizations often rely on visual cues, how do animals coordinate travel with distant, out-of-sight group members? The problem is particularly acute for animals like chimpanzees, bonobos, and spider monkeys that live in fission-fusion societies and inhabit dense forests where visibility is limited. In these species, members of a single social group regularly divide into smaller subgroups (or “parties”) that forage out-of-sight of one another, but frequently reunite after producing long-distance vocalizations (review across

taxa: Aureli et al. 2008; spider monkeys: Ramos-Fernandez 2005; Spehar and Di Fiore 2013; chimpanzees: Fedurek et al. 2014).

Long-distance calls are often thought to coordinate group travel because the calls allow individuals to locate and identify one another (e.g., elephants: Leighty et al. 2008; dolphins [*Stenella longirostris*]: Lammers et al. 2006; chimpanzees: Mitani and Nishida 1993; orangutans [*Pongo abelii*]: Mitra and van Schaik 2007). In some circumstances, however, both caller and receivers may benefit from signals that provide more information about a caller than only its identity and location. Chimpanzees, for example, produce food calls that vary according to the size of the tree in which the caller is feeding (Kalan et al. 2015). This information may benefit listeners by allowing them to assess the quality of distant feeding trees, while also aiding callers by more effectively recruiting preferred social partners. Similarly, hyenas (*Crocuta crocuta*) give long-distance “whoops” during conflicts with lions that differ from “display” whoops given in the absence of any conflict (Gersick et al. 2015). These acoustic differences permit listeners to distinguish low-urgency calls that merely signal the identity and location of other group members from high-urgency calls that require an immediate response to mobilize an effective mobbing response against lions.

Thus, animals that frequently separate and reunite may be under selection pressure to produce long-distance calls that can convey information about callers’ context, behavior, and motivation through acoustic cues alone, just as short-range travel-associated calls do in conjunction with visual cues.

In this study, we explore the role that long-distance calls play in coordinating the movement of distant individuals within a community of wild bonobos. Bonobos form long-term, stable communities, in which all members regularly associate with each other and share a home range (Kano 1992). Within a community, individuals form temporary parties that travel and forage separately from other parties. Parties are unpredictable in size (ranging from one individual to the entire community), duration (lasting from several minutes to several days), and composition.

Movement between parties is often preceded by loud calls, termed “high hoots” (HHs) (Hohmann and Fruth 1994; White et al. 2015). High hoots are tonal calls, given in a variety of non-aggressive contexts (de Waal 1988; Bermejo and Omedes 1999). They occur in bouts consisting of 1–27 acoustic units each with an inverted-U-shaped frequency contour (Hohmann and Fruth 1994) and are audible for at least 700 m in the forest (personal observation).

Under certain circumstances, bonobos combine HHs with other call types, including “whistle” (W) and “low hoot” (LH) vocalizations. Such call combinations may allow callers to signal their motivation to behave in a particular way more precisely than the production of single calls alone (Schlenker

et al. 2016; Seyfarth and Cheney 2016b). Indeed, in a previous paper, we investigated whistle-HH combinations (W+HHs) and found that individuals who produced W+HHs were more likely to approach another party than individuals who produced HHs alone, suggesting that production of the call combination was more informative about callers' subsequent behavior than a single call type (Schamberg et al. 2016).

Here, we examine bonobos' use of HHs in combination with LHs, an acoustically noisy, low-pitched vocalization in which the caller produces sound through both inspirations and expirations (de Waal 1988; Bermejo and Omedes 1999). Our behavioral observations indicate that LH+HH combinations often occur prior to or during travel. Furthermore, after producing LH+HH combinations, callers are likely to be approached by individuals from other parties. These two features suggest LHs+HHs may function to coordinate travel between separated individuals, but they do not explain how such coordination may occur. By examining caller behavior and comparing LH+HH combinations to both W+HH combinations and HHs alone, we consider the following four hypotheses about how LH+HH combinations function to coordinate inter-party movement:

1. *Contact call hypothesis.* Callers produce LHs+HHs to establish or maintain vocal contact with out-of-sight individuals. The calls facilitate reunions between separated individuals by providing listeners with information about the caller's identity and location. This hypothesis predicts that recruitments will be equally likely after LHs+HHs and HHs alone because both signals contain the same information about caller identity and location.
2. *Travel initiation hypothesis.* Callers produce LHs+HHs because they are motivated to initiate travel. Individuals from other parties are not necessarily the intended audience for these vocalizations, but they may nonetheless hear these vocalizations and decide to approach callers. This hypothesis predicts changes in the caller's party composition will not affect subsequent production of LHs+HHs because inter-party recruitment is not the caller's goal.
3. *Reunion hypothesis.* Callers produce both LHs+HHs and W+HHs when motivated to reunite with separated parties. Both combinations facilitate reunion by providing listeners with the same information about callers' motivation. This hypothesis predicts that there will be no difference in the *type* of reunion following each call combination—i.e., callers will be equally likely to approach another party as they are to recruit individuals to their own party, regardless of whether they produce LHs+HHs or W+HHs.
4. *Travel recruitment hypothesis:* Callers produce LHs+HHs to recruit individuals from other parties to travel with the caller's own party. Two predictions follow from this

hypothesis: (a) calls will be directed outside the caller's own party and (b) callers should cease calling once extra-party individuals have been successfully recruited.

## Methods

### Study site and subjects

For 13 months between July 2011 and March 2014, we sampled behavior and recorded vocalizations from 18 free-ranging adults (7 males and 11 females) in the Bompusa community at the LuiKotale field site in the Mai-Ndombe province of the Democratic Republic of Congo. Individuals in this community have been studied continuously since 2007 and were fully habituated and identified at the beginning of the study. The subjects' home range was approximately 40 km<sup>2</sup> and located in dense rainforest consisting of large patches of both terra firma and swamp forest (Hohmann and Fruth 2003).

### Data collection

Small groups of observers followed subjects on foot and collected data during half-day shifts (from either 0600 to 1200 or 1200 to 1800). Observers recorded behavioral observations with an audio recorder and later transcribed the data. Because the central research questions of this study did not depend on the rates of particular vocalizations or behaviors, observers primarily collected ad libitum data, allowing observers to sample key vocalizations and behaviors whenever they occurred (Altmann 1974). One thousand two hundred twenty-four hours of ad libitum data were collected.

When a subject produced a vocalization observers noted the following information:

1. *Identity of the caller.*
2. *Identities of other individuals in the caller's party.*
3. *Caller's location relative to the party.* Observers noted whether a caller was central or peripheral, relative to other individuals in the party. A caller was considered to be on the periphery of a party if, after having been engaged in a joint activity with members of its party, it increased its distance to 15–40 m from the majority of the party. For example, an individual who exited a tree before other members of the party and then vocalized on the ground while the other members of the party remained in the tree was considered to be on the periphery of the party. Similarly, if a party began to travel after leaving a feeding tree, but one individual remained in the tree and vocalized as the other members of the party were moving away, the caller was considered to be on the periphery of the party.

4. *Call type.* Observers provisionally noted the call type in the field, but final categorization was determined by listening to the recording and viewing the spectrogram (Fig. 1).
5. *Call context.* Observers noted whether the caller was resting, feeding, or traveling when vocalizing. If the caller was engaged in more than one of these activities, the context was considered ambiguous.
6. *Vocalizations from other parties.* Observers categorized each vocalization produced by callers as a “spontaneous” or “response” call. Spontaneous calls were those given in the absence of any calls by individuals outside the caller’s party during the 30 s prior to the call. Response calls were those produced within 30 s of vocalizations from another party. Observers also noted whether each call received a response—that is, was followed within 10 s by vocalizations from bonobos outside the caller’s party. We chose 10 s as the window for response vocalizations because bonobos occasionally take several seconds to respond to vocalizations. In most cases, however, responses were produced immediately after the prior vocalization. A call was considered to be part of a “vocal exchange” if it was a spontaneous call that received a response from another party, or if it was given in apparent response to call from another party.

For 5 min following a vocalization, observers continued to record the caller’s behavior and noted any changes—e.g., if the caller was feeding at the time of the vocalization, but then began to travel. Observers also noted the following changes in the composition of the caller’s party within 15 min after the production of a call: (1) no change, (2) one or more individual (s) left the caller’s party, (3) the caller approached and joined another party that was stationary, (4) one or more individual (s) from another party approached the caller’s party while the caller’s party was stationary, and (5) the caller’s party and one or more individual (s) from another party met while both parties were traveling. In our analysis, we termed the occurrence of (3) an “approach” and the occurrence of either (4) or (5) a “recruitment.”

In addition to the 1224 h of ad libitum data, observers obtained 117 h of data from all occurrence focal animal sampling (mean  $\pm$  SD  $6.6 \pm 1.7$  h/individual) (Altmann 1974). Each focal sample was 15 min, during which the observer followed a single subject and collected continuous data on the vocalizations and behaviors described above. Each individual in a party was sampled once before any individual was re-sampled. Individuals were never sampled more than once an hour. It was not possible for observers to be blind to the hypotheses being examined because the study was designed by the observers.

Observers also recorded party composition every 15 min. Party composition was defined as all individuals visible to observers or known to be within a radius of 50 m of the focal animal (Lehmann and Boesch 2004). Observers scanned the surrounding area and conferred with other observers in order to identify all animals in the party. Scans also included currently out-of-sight bonobos that were known to be present based on observations during the previous 15 min. These fixed-time party composition scans allowed us to calculate changes in party composition.

In addition to behavioral data, observers made continuous audio recordings of all vocalizations for which the caller and context could accurately be noted. Recordings were made using a PMD660 Marantz digital recorder and a Sennheiser ME66 microphone at sample rate of either 44.1 or 48 kHz. Spectrograms of audio files were created with WaveSurfer (version 1.88p) and RavenPro (version 1.5). Call types and call combinations were visually distinguishable in spectrograms (Fig. 1). Classification of call types followed descriptions of bonobo vocalizations in captivity (de Waal 1988) and the wild (Bermejo and Omedes 1999; Schamberg et al. 2016).

Observers also collected data on subjects’ ranging patterns on a GPS device (Garmin Map 62). Routes were created using the “tracks” function, which took a GPS point every minute.

### Data analysis

To answer our different questions, we used overlapping but non-identical datasets recorded during the same period. All results that report on the rates of behavior are based on focal samples; all other analyses pool ad libitum and focal data.

To analyze the context of call production, we included 65 observed LH+HH combinations and 135 HHs for which the context of the call and context in the 5 min after the call were known.

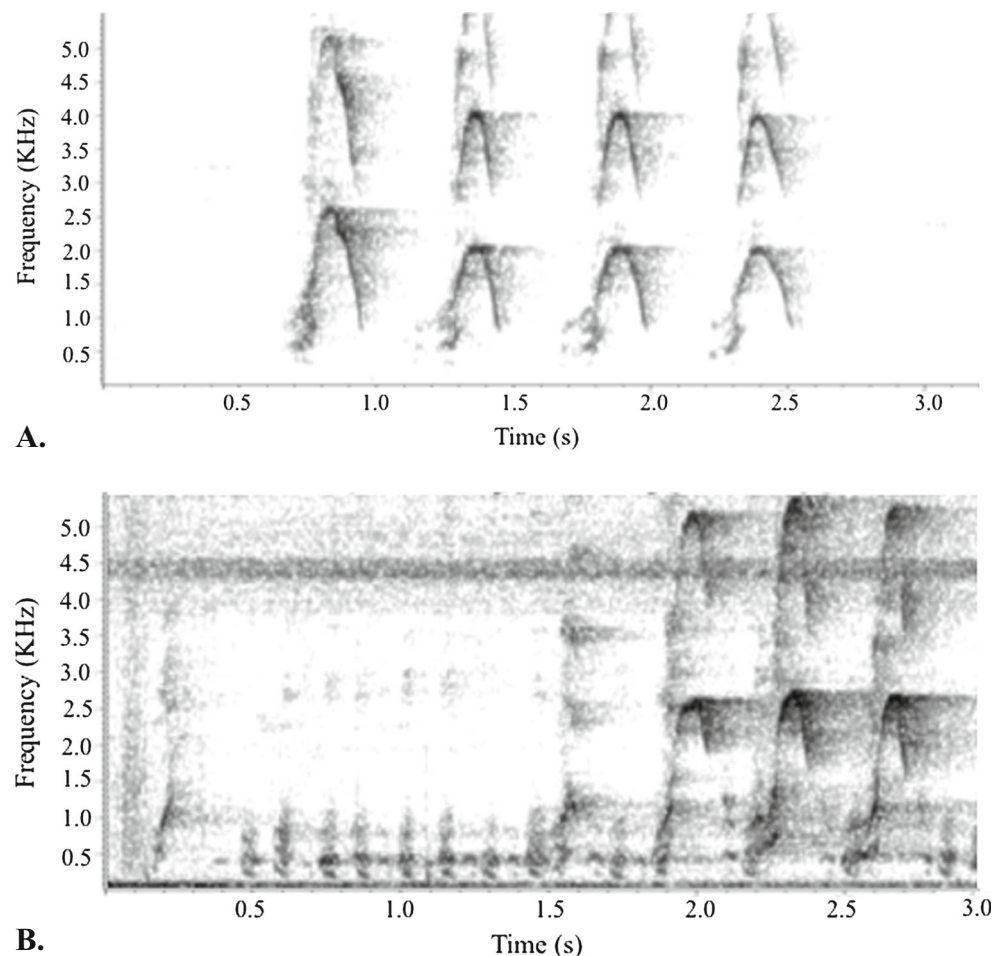
To examine changes in party composition after a call, we used observations of 40 LHs+HHs and 75 HHs for which we had reliable data on post-call changes in party composition (including observations when there was no change). In some cases, callers produced multiple HHs or LHs+HHs during a single communicative event. However, almost all changes of party composition in our analyses (93%) occurred after the production of the final calls in such a series. For the purposes of statistical independence, our analysis of changes in party composition following the production of different call types included only these final calls.

To calculate baseline rates of approaches and recruitments, we tabulated how many approaches and recruitments occurred during 468 all occurrence focal animal samples.

To test hypotheses about persistence in call production, we included all 54 observed LH+HH combinations for which we had reliable data on subsequent call production and any changes in party composition.



**Fig. 1** Spectrograms of **a** high hoots from an adult female and **b** a low hoot-high-high hoot combination from an adult male



## Statistical tests

To examine whether certain calls and call combinations were followed by different behaviors by callers, listeners, or both, we used Generalized Mixed Models (“glmer” function in “lmerTest” package in R version 3.1.2 GUI 1.65 Snow Leopard build (6833)). Because different individual callers contributed in different proportions to our pooled data, we entered caller ID as a random factor. Details of each model are provided in the [supplementary materials](#).

## Results

### 1. Call production

Bonobos produced bouts composed exclusively of HHs at an overall mean  $\pm$  SD rate of  $1.09 \pm 0.46$  calls per individual per hour. Individuals most commonly produced LHs in combination with other call types (HHs and “whistles”) and non-vocal signals (buttress drums and branch drag displays), but also produced LHs independently (Table 1). Individual bonobos produced signal combinations that contained LHs at an overall

mean  $\pm$  SD rate of  $0.15 \pm 0.19$  calls per hour per individual. However, call production was highly skewed. The first-ranking male produced 20% (21/104) of all combinations containing LHs, and the second-ranking male produced 53% (55/104) of all such combinations. Despite this skewed distribution, all males (7/7) and 45% (5/11) of all females were observed to produce at least one LH calling bout.

Table 1 lists the number of different signal combinations that included LHs. Because LHs were most commonly combined only with HHs (63% of all observed combinations), and

**Table 1** Total observations for each type of LH combination (data pooled from 12 subjects)

Signal	Observed cases
LH alone	4
LH+HH	65
LH+D	7
LH+BD	6
LH+HH+D	8
W+LH+HH	9
Total	99

LH low hoot, HH high hoot, W whistle, D buttress drum, BD branch drag

given the very small sample sizes for most of the other combination types, we examine only LH+HH combinations here. LH+HH combinations were not produced in any specific order: in 54% (35/65) of observed combinations, LHs preceded HHs, while in 46% (30/65) of instances, LHs followed HHs. We use “LHs+HHs” to refer to all combinations comprised of LHs and HHs, regardless of call order.

## 2. Call context

LHs+HHs were more likely than HHs alone to be produced during travel: 32% (21/65) of LHs+HHs were given while traveling, compared with 18% (24/134) of HHs alone (GLM  $\beta = 0.8$ ,  $SE = 0.3$ ,  $z = 2.2$ ,  $p = 0.025$ ). Additionally, LHs+HHs were more likely to occur prior to the initiation of travel than HHs alone. Forty-four percent (16/36) of LHs+HHs that were produced in a stationary context (i.e., while feeding or resting) were followed by travel in the subsequent 5 min, compared to 22% (17/85) of HHs alone (GLMM:  $\beta = 1.1$ ,  $SE = 0.4$ ,  $z = 3.6$ ,  $p = 0.009$ ).

Bonobos also traveled significantly faster after, as compared with before, production of the initial LHs+HHs. Based on 25 days when LHs+HHs were observed and GPS data were available, the mean  $\pm$  SD rate of travel before the production of the first LH+HH combinations was  $0.56 \pm 0.29$  km/h, compared to  $0.69 \pm 0.23$  km/h afterwards (paired  $t$  test:  $t = -2.8$ ,  $p = 0.01$ ).

## 3. Inter-party recruitment

In addition to their association with subsequent travel, LHs+HHs were significantly more likely than HHs alone to be associated with subsequent recruitments of individuals from other parties (GLMM:  $\beta_{GLMM} = 0.5$ ,  $SE = 0.5$ ,  $z = 3.3$ ,  $p = 0.001$ ) (Table 2). In contrast, callers were not more likely to approach another party after producing LHs+HHs than after producing HHs alone (GLMM:  $\beta = -0.4$ ,  $SE = 0.6$ ,  $z = -0.6$ ,  $p = 0.577$ ) (Table 2).

When callers produced LHs+HHs that did not result in recruitment, they produced additional LHs+HHs in 33% (13/39) of cases. In contrast, when callers produced LHs+HHs that did result in recruitment, they produced additional LHs+HHs in only 6% (1/16) of cases. The association between

LH+HH combinations that did not result in subsequent recruitments and callers producing additional calls approached significance (GLMM:  $\beta = -2.0$ ,  $SE = 1.1$ ,  $z = -1.8$ ,  $p = 0.064$ ).

Additionally, callers who produced LHs+HHs were on the periphery of their party in 28% (18/65) of cases, compared to 10% (14/134) of cases for HHs alone (GLMM:  $\beta = 1.0$ ,  $SE = 0.4$ ,  $z = 2.7$ ,  $p = 0.004$ ), possibly suggesting that the intended audience for LHs+HHs was not the caller's own party.

## 4. Comparison between LHs+HHs and W+HHs

In a previous paper, we showed that bonobos often produce high hoots in combination with a whistle (W+HHs) prior to approaching another party (Schamberg et al. 2016). In contrast, callers appeared to produce LHs+HHs prior to recruiting others to join their own party (Table 2). Comparing the two call types directly, we found that W+HHs were more likely to result in approaches than were LHs+HHs (GLMM:  $\beta = 2.3$ ,  $SE = 0.6$ ,  $z = 3.8$ ,  $p < 0.0001$ ). Conversely, LHs+HHs were more likely than W+HHs to result in the recruitment of others (GLMM:  $\beta = -1.7$ ,  $SE = 0.6$ ,  $z = -3.1$ ,  $p = 0.002$ ).

Previous research on W+HHs also showed that post-call behavior partially depended on whether the call was part of a vocal exchange: callers who produced W+HH combinations that were part of a vocal exchange were more likely to approach and join another party than callers who produced the same calls in the absence of a vocal exchange (Schamberg et al. 2016). Based on this finding, we examined the effect that vocal exchanges had on the outcome of LHs+HHs. The association between vocal exchanges and subsequent recruitments was not significant (GLMM:  $\beta = 0.1$ ,  $SE = 0.7$ ,  $z = 0.1$ ,  $p = 0.927$ ). Thus, LHs+HHs that were part of a vocal exchange were equally likely to result in recruitment as LHs+HHs that were not part of a vocal exchange.

## 5. Summary of differences between HHs, LHs+HHs, and W+HHs

We used odds ratios to compare the effects of HHs, LH+HHs, and W+HHs on subsequent approaches and recruitments (Table 3). Results show that each signal was associated with a different outcome: HHs increased the likelihood of both recruitments and approaches compared to baseline rates. W+HHs rarely led to recruitments, but sharply increased the likelihood of approaches if they were part of a call exchange. Finally, LHs+HHs rarely led to approaches, but increased the likelihood of a recruitment regardless of whether or not they were part of a call exchange or not.

**Table 2** Summary of post-call outcomes for HHs, LHs+HHs, and W+HHs

Call type	Approaches	Recruitments	Total observations
HHs	16	9	75
LHs+HHs	5	16	40
W+HHs	30	5	50

**Table 3** The likelihood of either “approaches” or “recruitments” after specific calls or call combinations, compared to baseline rates. The numbers in each cell represent the odds ratio, calculated as the odds of a specific outcome following a specific call type divided by the odds of that outcome during baseline observations (Tabachnick and Fidell 2007:462). Baseline data were derived from focal animal samples. Data on W+HHs are taken from Schamberg et al. (2016)

Call type	Recruitment	Approach
HH	2.6	3.4
W+HH (no call exchange)	2.2	8.2
W+HH (call exchange)	0.9	34.0
LH+HH (no call exchange)	11.9	0.0
LH+HH (call exchange)	12.7	3.1

## Discussion

Bonobos used a specific call combination (LHs+HHs) during inter-party travel recruitment. LHs+HHs were more likely than a single call type alone (HHs) to be given by callers who were traveling. Similarly, when produced while resting or feeding, LHs+HHs were more likely than HHs alone to signal imminent travel. Bonobos also traveled at a faster rate after producing LH+HH combinations. In addition to their general association with travel, LHs+HHs were more likely than HHs alone or the “whistle-high hoot” combination (W+HHs) to result in the recruitment of extra-party individual(s).

These results cannot be explained by the *Contact call*, *Travel initiation*, or *Reunion hypotheses*. The *Contact call hypothesis* held that both HHs alone and LH+HH combinations contained information only about a caller’s identity and location, and, therefore, predicted similar outcomes. This prediction is not consistent with our results, which demonstrated a clear difference in the behavior of both callers and listeners following the production of LHs+HHs as opposed to HHs alone. The *Travel initiation hypothesis*, which posited that LHs+HHs were produced with the goal of spurring group travel, predicted that successful recruitments would not affect call production. Contrary to this prediction, callers tended to stop producing LHs+HHs after successful recruitments. Finally, under the *Reunion hypothesis*, callers produce both LHs+HHs and W+HHs when motivated to reunite with separated parties, and we should observe no difference in the *type* of reunion following each call combination. Post-call outcomes for the two call combinations, however, indicated that LHs+HHs were much more likely to be used to recruit extra-party individuals, while W+HHs were much more likely to result in the caller approaching and joining another party.

In contrast, our results were consistent with the *Travel recruitment hypothesis*, which held that callers produce LHs+HHs to recruit individuals from other parties to travel with caller’s own party. This hypothesis predicted that callers would direct LHs+HHs outside the caller’s own party, and

that caller’s would continue to call until they successfully recruited extra-party individuals to their own party. Our data supported both predictions: callers who gave LHs+HHs were more likely than others to be on the periphery of their own party, and callers tended to continue to produce LHs+HHs until they were approached by individuals from another party.

The use of LHs+HHs to recruit extra-party individuals is particularly intriguing when considered alongside HHs and W+HHs. Bonobos’ long-distance call repertoire seems to consist of a general-purpose long-distance call (HHs) that, under certain circumstances, is combined with other call types (W and LHs) to convey information beyond caller identity and location. Specifically, W+HHs appear to signal a caller’s motivation to approach and join another party, while LHs+HHs signal the reverse situation in which a caller seeks to recruit extra-party individuals.

HHs alone undoubtedly play an important role in inter-party communication and travel coordination (Hohmann and Fruth 1994; White et al. 2015). However, given their usage across various contexts, it is likely that, in the absence of additional contextual cues, they only provide listeners with ambiguous information about a caller’s motivation and likely subsequent behavior. By combining HHs with LHs and Ws, bonobos provide listeners with more precise information about caller’s motivation and likely behavior. In the unpredictable context of fission-fusion travel coordination—will A move to B or vice versa?—calls that reduce uncertainty may benefit both callers and receivers by facilitating reunions that are necessary to maintain social relationships and gain access to mates (Sueur et al. 2011).

The use of vocalizations to facilitate travel coordination and behavioral synchronization is not, of course, unique to fission-fusion societies (Boinski and Garber 2000). In species with more cohesive social structures, however, single call types that are produced in a variety of contexts may suffice to synchronize group movement because additional visual cues allow listeners to correctly interpret such underspecified calls (Kondo and Watanabe 2009). Communication between separated individuals may demand that signals convey more specific information through acoustic cues alone. Thus, the use of LH+HH and W+HH combinations to coordinate inter-party movement may have evolved in response to the demands a low-visibility fission-fusion social structure.

## Conclusion

Bonobos use two distinct call combinations in the context of inter-party travel coordination. The combinations provide listeners with qualitatively different information about a caller’s motivation to either approach another party, or recruit other individuals to the caller’s own party. The use of call combinations in this context may have been favored by natural

selection because of limitations inherent in long-distance communication. In the absence of visual cues, calls that are able to provide information through acoustic cues alone about a caller's motivation and likely behavior may benefit both callers and receivers.

**Acknowledgments** We thank the Institut Congolais pour la Conservation de la Nature (ICCN) for permission to conduct research in the Democratic Republic of Congo. We are grateful to Lys Alcayna for data collection in the field and three anonymous reviewers for comments and suggestions.

#### Compliance with ethical standards

**Ethical approval** The study was conducted in accordance with the current laws in the USA, Germany, and the Democratic Republic of the Congo. Research was approved by the Animal Care and Use Committee of the University of Pennsylvania (Protocol no. 804117).

**Conflict of interest** The authors declare that they have no conflicts of interest.

**Funding** This research was funded by the National Science Foundation Graduate Research Fellowship (IS) and grants from the Leaky Foundation (IS and ZC), the Department of Psychology at the University of Pennsylvania (IS), National Geographic Society (RMS and ZC) (Grant #9115-12).

#### References

- Altmann J (1974) Observational study of behavior: sampling methods. *Behaviour* 49:227–267
- Aureli F, Schaffner CM, Boesch C et al (2008) Fission-fusion dynamics: new research frameworks. *Curr Anthropol* 49:627–654
- Bermejo M, Omedes A (1999) Preliminary vocal repertoire and vocal communication of wild bonobos (*Pan paniscus*) at Lilungu (Democratic Republic of Congo). *Folia Primatol* 70:328–357
- Boinski S (1993) Vocal coordination of troop movement among white-faced capuchin monkeys, *Cebus capucinus*. *Am J Primatol* 30:85–100
- Boinski S, Garber PA (2000) On the move: how and why animals travel in groups. University of Chicago Press, Chicago
- Bousquet CAH, Sumpter DJT, Manser MB (2011) Moving calls: a vocal mechanism underlying quorum decisions in cohesive groups. *Proc R Soc Lond B* 278:1482–1488
- Clay Z, Archbold J, Zuberbühler (2015) Functional flexibility in wild bonobo vocal behaviour. *Peer J* 3:e1124
- da Cunha RGT, Byrne RW (2009) The use of vocal communication in keeping the spatial cohesion of groups: intentionality and specific functions. In: Garber PA, Estrada A, Bicca-Marques JC, Heymann EW, Strier KB (eds) *South American Primates*. Springer, New York, pp 341–363
- de Waal FBM (1988) The communicative repertoire of captive bonobos (*Pan paniscus*), compared to that of chimpanzees. *Behaviour* 106:183–251
- Fedurek P, Donnellan E, Slocombe KE (2014) Social and ecological correlates of long-distance pant hoot calls in male chimpanzees. *Behav Ecol Sociobiol* 68:1345–1355
- Fischer J, Zinner D (2011) Communication and cognition in primate group movement. *Int J Primatol* 32:1279–1295
- Gersick AS, Cheney DL, Schneider JM, Seyfarth RM, Holekamp KE (2015) Long-distance communication facilitates cooperation among wild spotted hyaenas, *Crocuta crocuta*. *Anim Behav* 103:107–116
- Gruber T, Zuberbühler K (2013) Vocal recruitment for joint travel in wild chimpanzees. *PLoS One* 8:e76073
- Hohmann G, Fruth B (1994) Structure and use of distance calls in wild bonobos (*Pan paniscus*). *Int J Primatol* 15:767–782
- Hohmann G, Fruth B (2003) Lui Kotal—a new site for field research on bonobos in the Salonga National Park. *Pan Africa News* 10:25–27
- Kalan AK, Mundry R, Boesch C (2015) Wild chimpanzees modify food call structure with respect to tree size for a particular fruit species. *Anim Behav* 101:1–9
- Kanō T (1992) The last ape: pygmy chimpanzee behavior and ecology. Stanford University Press, Stanford
- Kondo N, Watanabe S (2009) Contact calls: information and social function. *Jpn Psychol Res* 51:197–208
- Lammers MO, Schotten M, Au WW (2006) The spatial context of free-ranging Hawaiian spinner dolphins (*Stenella longirostris*) producing acoustic signals. *J Acoust Soc Am* 119:1244–1250
- Lehmann J, Boesch C (2004) To fission or to fusion: effects of community size on wild chimpanzee (*Pan troglodytes verus*) social organisation. *Behav Ecol Sociobiol* 56:207–216
- Leighty KA, Soltis J, Wesolek CM, Savage A (2008) Rumble vocalizations mediate interpartner distance in African elephants, *Loxodonta africana*. *Anim Behav* 76:1601–1608
- Mitani JC, Nishida T (1993) Contexts and social correlates of long-distance calling by male chimpanzees. *Anim Behav* 45:735–746
- Mitra S, van Schaik CP (2007) The response of adult orangutans to flanged male long calls: inferences about their function. *Folia Primatol* 78:215–226
- Poole JH (2011) Behavioral contexts of elephant acoustic communication. In: Moss CJ, Croze HJ, Less PC (eds) *The Amboseli elephants: a long-term perspective on a long-lived mammal*. University of Chicago Press, Chicago, pp 125–161
- Price T, Wadewitz P, Cheney DL, Seyfarth RM, Hammerschmidt K, Fischer J (2015) Vervets revisited: a quantitative analysis of alarm call structure and context specificity. *Sci Rep* 5:13220
- Ramos-Fernández G (2005) Vocal communication in a fission–fusion society: do spider monkeys stay in touch with close associates? *Int J Primatol* 26:1077–1092
- Rendall D, Seyfarth RM, Cheney DL, Owren MJ (1999) The meaning and function of grunt variants in baboons. *Anim Behav* 57:583–592
- Schamberg I, Cheney DL, Clay Z, Hohmann G, Seyfarth RM (2016) Call combinations, vocal exchanges and interparty movement in wild bonobos. *Anim Behav* 122:109–116
- Schlenker P, Chemla E, Schel AM, Fuller J, Gautier JP, Kuhn J, Veselinović D, Arnold K, Cäsar C, Keenan S, Lemasson A, Ouattara K, Ryder R, Zuberbühler K (2016) Formal monkey linguistics. *Theoretical Linguistics* 42:1–90
- Seyfarth RM, Cheney DL (2016a) The origin of meaning in animal signals. *Anim Behav* 124:339–346
- Seyfarth RM, Cheney DL (2016b) Schlenker et al.'s informativity principle. *Theor Linguist* 42:155–158
- Spehar SN, Di Fiore A (2013) Loud calls as a mechanism of social coordination in a fission–fusion taxon, the white-bellied spider monkey (*Ateles belzebuth*). *Behav Ecol Sociobiol* 67:947–961
- Stewart KJ, Harcourt AH (1994) Gorillas' vocalizations during rest periods: signals of impending departure? *Behaviour* 130:29–40
- Sueur C, King AJ, Conradt L, Kerth G, Lusseau D, Mettke-Hofmann C, Schaffner CM, Williams L, Zinner D, Aureli F (2011) Collective decision-making and fission–fusion dynamics: a conceptual framework. *Oikos* 120:1608–1617
- Tabachnick BG, Fidell LS (2007) Using multivariate statistics. Pearson Education, Boston
- Wheeler BC, Fischer J (2012) Functionally referential signals: a promising paradigm whose time has passed. *Evol Anthropol* 21:195–205
- White FJ, Waller M, Boose K, Merrill MY, Wood KD (2015) Function of loud calls in wild bonobos. *J Anthropol Sci* 93:1–13