Title
Correction for Bryant et al., Detecting affiliation in colaughter across 24 societies

Journal
Proceedings of the National Academy of Sciences of the United States of America, 113(21)

ISSN
0027-8424

Authors
Bryant, Gregory A
Fessler, Daniel MT
Fusaroli, Riccardo
et al.

Publication Date
2016-05-24

DOI
10.1073/pnas.1606204113

Peer reviewed

Detecting affiliation in co-laughter across 24 societies

Gregory A. Bryant [1,2]*, Daniel M. T. Fessler [2,3], Riccardo Fusaroli [4,5], Edward Clint [2,3], Lene Aarøe [6], Coren L. Apicella [7], Michael Bang Petersen [8], Shaneikiah T. Bickham [9], Alexander H. Bolyanatz [10], Brenda Chavez [11], Delphine De Smet [12], Cinthya Díaz [11], Jana Fančovičová [13], Michal Fux [14], Paulina Giraldo-Perez [15], Anning Hu [16], Shanmukh V. Kamble [17], Tatsuya Kameda [18], Norman P. Li [19], Francesca R. Luberti [2,3], Pavol Prokop [13], Katinka Quintelier [20], Brooke Scelza [2,3], Hyun Jung Shin [21], Montserrat Soler [22], Stefan Stieger [23,24], Wataru Toyokawa [25], Ellis A. van den Hende [26], Hugo Viciana-Asensio [27], Saliha Elif Yildizhan [28], Jose C. Yong [19], Tessa Yuditha [29], Yi Zhou [16]

[1] Department of Communication Studies, University of California, Los Angeles.

[2] UCLA Center for Behavior, Evolution, and Culture.

[3] Department of Anthropology, University of California, Los Angeles.

[4] Interacting Minds Center, Aarhus University.

[5] Center for Semiotics, Aarhus University.

[6] Department of Political Science and Government, Aarhus University.

[7] Department of Psychology, University of Pennsylvania.

[8] Department of Political Science and Government, Aarhus University.

[9] The American International School in Egypt.

[10] College of DuPage.

[11] Pontificia Universidad Catolica del Peru.

[12] Ghent University.

[13] Department of Biology, University of Trnava.

[14] Department of Biblical and Ancient Studies, University of South Africa.

[15] University of Auckland.

[16] Fudan University.

[17] Department of Psychology, Karnatak University Dharwad.

[18] Department of Social Psychology, University of Tokyo.

[19] Singapore Management University.

[20] University of Amsterdam.

[21] Pusan National University, Korea.

[22] Department of Anthropology, Montclair State University.

[23] Department of Psychology, University of Konstanz.

[24] School of Psychology, University of Vienna.

[25] Department of Behavioral Science, Hokkaido University.

[26] Delft University of Technology.

[27] Université Paris 1 Panthéon-Sorbonne.

[28] Istanbul Bilgi University.

[29] Jakarta Field Station, Max Planck Institute for Evolutionary Anthropology.

*Correspondence to: G.A. Bryant, Department of Communication Studies, University of California, Los Angeles, 2303 Rolfe Hall, Los Angeles, CA 90095-1563. Tel.: +1 310 825 5984. E-mail address: gabryant@ucla.edu.

Classification: Social (major); Psychological and Cognitive Sciences (minor)

CITATION: Proceedings of the National Academy of Sciences 113(21):E3051 24 May 2016

Abstract

Laughter is a nonverbal vocal expression that often communicates positive affect and cooperative intent in humans. Temporally coincident laughter occurring within groups is a potentially rich cue of affiliation to overhearers. We examined listeners' judgments of affiliation based on brief, decontextualized instances of co-laughter between either established friends or recently acquainted strangers. In a sample of 966 participants from 24 societies, people reliably distinguished friends from strangers with an accuracy of 53% - 67%. Acoustic analyses of the individual laughter segments revealed that, across cultures, listeners' judgments were consistently predicted by voicing dynamics, suggesting perceptual sensitivity to emotionally triggered spontaneous production. Co-laughter affords rapid and accurate appraisals of affiliation that transcend cultural and linguistic boundaries, and may constitute a universal means of signaling cooperative relationships.

Keywords: laughter, nonverbal behavior, cooperation

Significance Statement

Human cooperation requires reliable communication about social intentions and alliances. Though laughter is a phylogenetically conserved vocalization linked to affiliative behavior in nonhuman primates, its functions in modern humans are not well understood. We show that judges all around the world, hearing only brief instances of co-laughter produced by pairs of American English speakers in real conversations, are able to reliably identify friends and strangers. Participants' judgments of friendship status were linked to acoustic features of laughs known to be associated with spontaneous production and high arousal. These findings strongly suggest that co-laughter is universally perceivable as a reliable indicator of relationship quality, and contribute to our understanding of how nonverbal communicative behavior might have facilitated the evolution of cooperation.

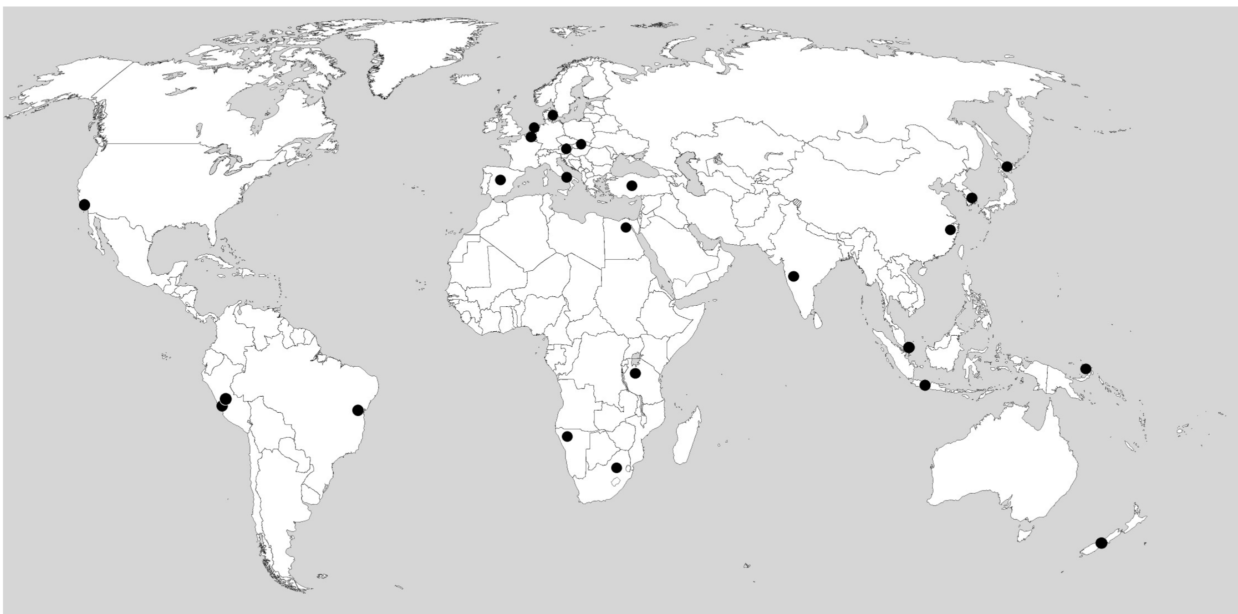
Humans exhibit extensive cooperation between unrelated individuals, managed behaviorally by a suite of elaborate communication systems. Social coordination relies heavily on language, but nonverbal behaviors also play a crucial role in forming and maintaining cooperative relationships (1). Laughter is a common nonverbal social vocalization that universally manifests across a broad range of contexts, and is often associated with prosocial intent and positive emotions (2-5). Despite the ubiquity and similarity of laughter across all cultures, its communicative functions remain largely unknown. Co-laughter is simultaneous laughter between individuals in social interactions, and occurs with varying frequency as a function of the sex and relationship composition of the group: friends laugh together more than strangers, and groups of female friends tend to laugh more than groups of male friends or mixed-sex groups (6, 7). Researchers have focused on laughter within groups, but co-laughter potentially provides rich social information to those outside of the group. Against this backdrop, we examined i) whether listeners around the world can determine the degree of social closeness and familiarity between pairs of people solely on the basis of very brief decontextualized recordings of co-laughter, and ii) which acoustic features in the laughs might influence such judgments.

Laughter is characterized by neuromechanical oscillations involving rhythmic laryngeal and superlaryngeal activity (8, 9). It often features a series of bursts or calls, collectively referred to as bouts. Laugh acoustics vary dramatically both between and within speakers across bouts (10), but laughter appears to follow a variety of production rules (11). Comparative acoustic analyses examining play vocalizations across several primate species suggest that human laughter is derived from a homolog dating back at least 20 MYA (12, 13). Humans evolved species-specific sound features in laughs involving higher proportions of periodic components (i.e., increasingly voiced), and a predominantly egressive airflow. This pattern is different from laugh-like vocalizations of our closest nonhuman relative, *Pan troglodytes*, which incorporate alternating airflow, and mostly noisy, aperiodic structure (2, 12). In humans, relatively greater voicing in laughs is judged to be more emotionally positive than unvoiced laughs (14), as is greater variability in pitch and loudness (15). People produce different perceivable laugh types (e.g., spontaneous [or Duchenne] versus volitional [or non-Duchenne]) that correspond to different communicative functions and underlying vocal production systems (3, 16-18), with spontaneous laughter produced by an emotional vocal system shared by many mammals (19, 20). Recent evidence suggests that spontaneous laughter is often associated with relatively greater arousal in production (e.g., increased pitch and loudness) than volitional laughter, and contains relatively more features in common with nonhuman animal vocalizations (16) (audio samples S1-S6 of different laughter types are available in the supplementary materials). These acoustic differences might be important for making social judgments if the presence of spontaneous (i.e., genuine) laughter predicts cooperative social affiliation, but the presence of volitional laughter does not.

All perceptual studies to date have examined individual laughs, but laughter typically occurs in social groups, often with multiple simultaneous laughers. Both because social dynamics can change rapidly and because newcomers will often need to quickly assess the membership and boundaries of coalitions, listeners frequently must make rapid judgments about the affiliative status obtaining within small groups of interacting individuals; laughter may provide an efficient and reliable cue of affiliation. If so, we should expect humans to exhibit perceptual adaptations sensitive to co-laughter dynamics between speakers. However, to date, no study has examined listeners' judgments of the degree of affiliation between laughers engaged in spontaneous social interactions.

We conducted a cross-cultural study examining listeners' judgments of co-laughter produced by dyads composed either of friends or newly-acquainted strangers, with listeners hearing only extremely brief decontextualized recordings of co-laughter. This "thin slice" approach is useful because listeners receive no extraneous information that could inform their judgments, and success with such limited information indicates particular sensitivity to the stimulus (21). A broadly cross-cultural sample is important if we are to demonstrate the independence of this perceptual sensitivity from the influences of language and culture (22). While cultural factors likely shape pragmatic considerations driving human laughter behavior, we expect that many aspects of this phylogenetically ancient behavior will transcend cultural differences between disparate societies.

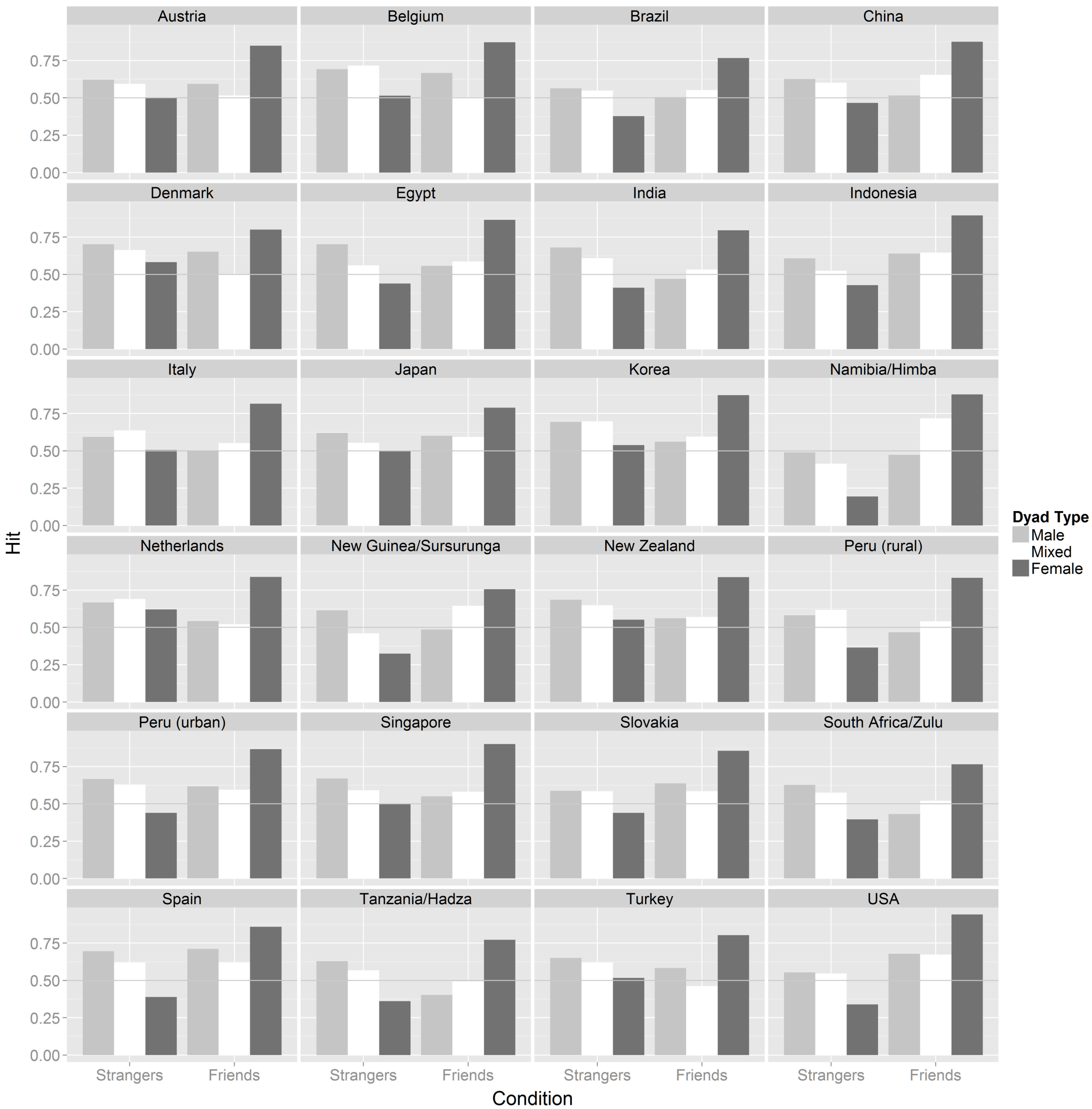
Fig. 1. Map of the 24 study site locations.



Results

Judgment task. We used a model comparison approach in which variables were entered into generalized linear mixed models (GLMMs) and effects on model fit were measured using Akaike Information Criterion (AIC) (see supplementary materials for details of all models). This approach allows researchers to assess which combination of variables best fit the pattern of data without comparison to a null model. The data were modeled using the glmer procedure of the lme4 package (23) in the statistical platform R (version 3.1.1) (24). Our dependent measures consisted of two questions: one forced-choice item and one rating scale. For Question 1 (*Were the speakers friends or strangers at the time of interaction?*) data were modeled using a binomial (logistic) link function, with judgment accuracy (hit rate) as a binary outcome (1 = correct; 0 = incorrect). For Question 2 (*How much did the speakers like one another?*), we used a Gaussian link function with rating response (1-7) as a continuous function.

Fig. 2. Rates of correct judgments (hits) in each study site broken down by experimental condition (friends or strangers), and dyad type (male-male, male-female, female-female). Chance performance represented by 0.50.



Across all participants, the overall rate of correct judgments was 61% (*SD* = 0.49), a performance significantly better than chance ($z = 40.5$, $p < 0.0001$). Figure 2 shows the rates of correct judgments for each study site with means and standard deviation values in Table S7. For the forced-choice measure (friends or strangers), the best-fitting model was a GLMM by the Laplace approximation, with participant sex as a fixed effect, familiarity and dyad type as interacting fixed effects, participant and study site as random effects, and hit rate (% correct) as the dependent measure; see Table 1. Participants (*VAR* = 0.014; *SD* = 0.12) and study site (*VAR* = 0.028; *SD* = 0.17) accounted for very little variance in accuracy in the forced-choice measure. Familiarity interacted with dyad type with female friends being recognized at higher rates than male friends (z = 42.96, P < 0.001), while male strangers were recognized at higher rates than female strangers (z = -22.57, P < 0.0001). A second significant interaction indicating that mixed-sex friends were recognized at higher rates than male friends, and mixed-sex strangers were recognized at lower rates than male strangers (z = 4.42, P < 0.001). For the second question (i.e., "How well do these people like each other?") the same model structure was the best fit, with a similar pattern of results. Question 2 results are displayed in Figure S1, with means and standard deviation values in Table S8

Table 1. Best-fit model for Question 1.

Random effects			**Fixed effects**				
Factor	***Variance***	***STD***	***Factor***	***Estimate***	***SE***	***z***	***Pr(>|z|)***
Subject	0.01469	0.1212					
Society	0.02772	0.1649					
			Condition	-0.31151	0.03818	-8.16	3.39e-16 ***
			Sex	0.05295	0.02226	2.38	0.017384 *
			ConvType1	-0.13943	0.03604	-3.87	0.000109 ***
			ConvType2	-0.75764	0.03436	-22.05	< 2e-16 ***
			Condition × ConvType1	0.19383	0.05038	3.85	0.000119 ***
			Condition × ConvType2	2.13094	0.05130	41.54	< 2e-16 ***

Note: ***: p < 0.001; *: p < 0.05.

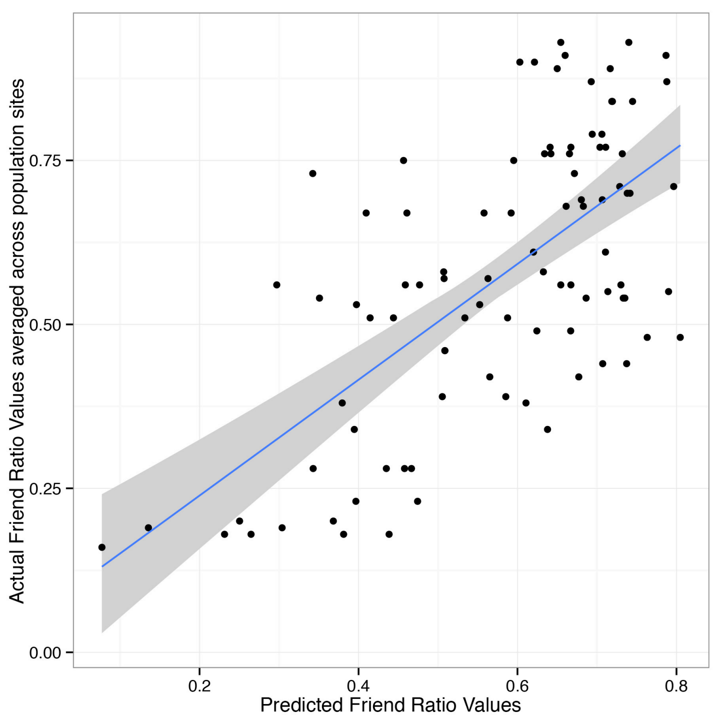
Overall, female friends were identified at the highest rate in every society without exception, but there was also a universal tendency to judge female co-laughers as friends (See Fig. S2). Forced-choice responses for each co-laughter trial were collapsed across societies and compared across dyad types revealing that a response bias to answer "friends" existed in judgments of female dyads (70%), F(2, 47) = 7.25, P = 0.002, but not male (46%) or mixed-sex dyads (49%), which did not differ from one another (LSD Test, P = 0.73). Additionally, female participants (M = 0.62; SD = 0.49) had slightly greater accuracy than male participants (M = 0.60; SD = 0.49) overall (z = 2.31, P < 0.05).

Acoustic analysis. Acoustic features, including the frequency and temporal dynamics of voiced and unvoiced segments, were automatically extracted from the individual laugh segments and employed to statistically reconstruct the rate at which participants judged each co-laugh segment as a friendship dyad. We employed an ElasticNet process (25) to individuate key features to assess in a multiple linear regression and a 5-fold cross-validated multiple linear regression to estimate coefficients of the selected features, repeating the process 100 times to ensure stability of the results (see Table 2). The resulting model was able to reliably predict participants' judgments that co-laughers were friends, adjusted $R^2 = 0.43$ (CI: 0.42-0.43), $P = 0.0001$ (Fig. 3). Across cultures, laughs that had shorter call duration, higher pitch and intensity irregularity, together with less variable levels of pitch irregularity were more likely to be judged to be between friends (for complete details of acoustic analysis see SI text).

Table 2. Sample coefficients from one run of the 5-fold cross-validated model for Friend Ratio across 24 societies.

Predictor	Beta (SE) Fold 1	Beta (SE) Fold 2	Beta (SE) Fold 3	Beta (SE) Fold 4	Beta (SE) Fold 5
Intercept	0.611 (0.177)	0.578 (0.114)	0.547 (0.114)	0.566 (0.125)	0.594 (0.104)
Jitter mean	1.720 (0.345)	1.652 (0.306)	1.648 (0.328)	1.545 (0.335)	1.720 (0.300)
Jitter SD	-1.826 (0.325)	-1.797 (0.305)	-1.747 (0.302)	-1.697 (0.338)	-1.9 (0.297)
5th PCTL shimmer	0.280 (0.199)	0.290 (0.131)	0.315 (0.127)	0.324 (0.146)	0.302 (0.119)
Mean call duration	-0.387 (0.075)	-0.358 (0.08)	-0.37 (0.084)	-0.412 (0.09)	-0.385 (0.07)

Fig. 3. Acoustic-based model predictions of Friends Ratio (on the x-axis) with the actual values (on the y-axis) (95% CI).



Discussion

Across all societies, listeners were able to distinguish pairs of co-laughers who were friends from those who were strangers that had just met. Biases, presumably reflecting panhuman patterns in the occurrence of laughter in everyday life, existed in all societies sampled as well, such that participants were more likely to assume that female co-laughers were friends than strangers. Male strangers were also recognized universally at significantly high rates, and participants worldwide rated the members of these dyads as liking each other the least among all pairs. Dynamic acoustic information in the laughter predicted the accuracy of judgments, strongly suggesting that participants attended closely to these sound features, likely outside of conscious awareness. The judgment pattern was remarkably similar across disparate societies, including those with essentially no familiarity with English, the language of the target individuals whose laughter was evaluated. These results constitute strong preliminary evidence that co-laughter provides a reliable cue with which overhearers (and, presumably, co-laughers themselves) can assess the degree of affiliation between interactants. Though embedded within discourse, laughter is nonverbal in nature, and presents universally interpretable features, presumably reflecting phylogenetic antiquity predating the evolution of language.

Together with auxiliary experiments on the laugh stimuli described in the SI Appendix, acoustic data strongly suggest that individual laugh characteristics provided much of the information allowing our participants to accurately differentiate between friends and strangers. Laugh features predicting listeners' friend responses included shorter call duration, associated with judgments of friendliness (14) and spontaneity (16), as well as greater pitch and loudness irregularities, associated with speaker arousal (26). Acoustic analyses comparing laughs within a given co-pair did not indicate any contingent dynamic relationship that could plausibly correspond to percepts of entrainment or coordination one might expect from familiar interlocutors. Indeed, our co-laugh audio clips may be too short to capture shared temporal dynamics that longer recordings might reveal. A second group of U.S. listeners evaluated artificial co-laughter pairs constructed by shuffling the individual laugh clips within dyad categories (see SI Appendix). Consonant with the conclusion that our main result was driven by features of the individual laughs rather than interactions between them, these artificial co-pairs were judged quite similarly to the original co-pairs in the main study. Lastly, a third group of U.S. listeners rated the individual laughs on the affective dimensions of arousal and valence; these judgments were positively associated with the likelihood that, in its co-laughter context, a given laugh was judged in the main study as having occurred in a friendship dyad.

Given the centrality in social dynamics of cooperation among allied individuals, and the fluidity with which relationships can change, in our species' ancestral past, individuals who could accurately assess the degree of affiliation between others stood to gain substantial fitness benefits. Closely allied individuals often constitute formidable opponents; likewise, such groups may provide substantial benefits to newcomers who are able to gain entry. Many social primates exhibit these political dynamics, along with corresponding cognitive abilities; by virtue of the importance of cooperation in human social and economic activities, ours is arguably the political species par excellence. Yet, even as language and cultural evolution have provided avenues for evolutionarily unprecedented levels of cooperation and political complexity in humans, we continue to employ vocal signals of affiliation that apparently predate these innovations. This opens up a host of evolutionary questions concerning laughter. Can affiliative laughter be simulated effectively, or is it an unfakeable signal? Hangers-on might do well to deceptively signal to overhearers that they are allied with a powerful coalition, while others would benefit from detecting such deception. If the signal is indeed honest, what keeps it so? Does the signal derive from the relationship itself, i.e., can unfamiliar individuals allied due to expedience signal their affiliation through laughter, or, consonant with the importance of coordination in cooperation, is intimate knowledge of the other party a prerequisite? Paralleling such issues, at the proximate level, numerous questions remain. For example, given universal biases that apparently reflect prior beliefs, future studies should both explore listeners' accuracy in judging the sex of co-laughers and examine the sources of such biases. Our finding that co-laughter constitutes a panhuman cue of affiliation status is thus but the tip of the iceberg when it comes to understanding this ubiquitous yet understudied phenomenon.

Methods

Participants.

Stimuli. All co-laughter segments were extracted from conversation recordings, originally collected for a project unrelated to the current study, made in 2003 at the Fox Tree laboratory at the University of California, Santa Cruz. The recorded conversations were between pairs of undergraduate students who volunteered to participate in exchange for course credit for an introductory class in psychology. Two rounds of recruitment were held. In one, participants were asked to sign up with a friend whom they had known for any amount of time. In the other, participants were asked to sign up as individuals, where after they were paired with a stranger. The participants were instructed to talk about any topic of their choosing; "bad roommate experiences" was given as an example of a possible topic. The average length of the conversations from which the stimuli employed in this study were extracted was 13.5 minutes (mean length ± SD = 809.2 s. ± 151.3 s.). Interlocutors were recorded on separate audio channels using clip-on lapel microphones (Sony ECM-77B) placed approximately 15 cm from the mouth, and recorded to DAT (16-bit amplitude resolution, 44.1 kHz sampling rate, uncompressed wav files, Sony DTC series recorder). For more description of the conversations, see (27).

Co-laughter segments. 48 co-laughter segments were extracted from 24 conversations (2 from each), half from conversations between established friends (mean length of acquaintance = 20.5 mo.; range = 4-54 mo.; mean age ± SD = 18.6 ± 0.6) and half from conversations between strangers who had just met (mean age ± SD = 19.3 ± 1.8). Co-laughter was defined as the simultaneous vocal production (intensity onsets within 1 s), in two speakers, of a nonverbal, egressive, burst series (or single burst), either voiced (periodic) or unvoiced (aperiodic). Laughter is acoustically variable, but often stereotyped in form, characterized typically by an initial alerting component and a decay function (2, 8, 12). Voiced laughter also generally contains a set fundamental frequency (F_0) component and stable vowel configuration that decays over time (9).

In the co-laughter segments selected for use, no individual laugh contained verbal content or other noises of any kind. To prevent a selection bias in stimulus creation, for all conversations, only two co-laughter sequences were used, namely the first to appear in the conversation, and the last to appear in the conversation. If a co-laughter sequence identified using this rule contained speech or other noises, the next qualifying occurrence was chosen. The length of co-laughter segments (in ms.) between friends (mean length ± SD = 1146 ± 455) and strangers (mean length

± SD = 1067 ± 266) was similar, $t(46) = 0.74$, $P = 0.47$. Laughter onset asynchrony (in ms.) was also similar between friends (mean length ± SD = 337 ± 299) and strangers (mean length ± SD = 290 ± 209), $t(46) = -0.64$. $P = 0.53$. Previous studies have documented that the frequency of co-laughter varies as a function of the gender composition of the dyad or group (6, 11). The same was true in the source conversations employed here, with female friends producing co-laughter at the highest frequency, followed by mixed-sex groups, and then all-male groups. Consequently, our stimulus set had uneven absolute numbers of different dyad types, reflecting the actual occurrence frequency in the sample population. Of the 24 sampled conversations, 10 pairs were female dyads, 8 pairs were mixed-sex dyads, and 6 pairs were male dyads. For each of these sex pair combinations, half were friends, and half were strangers.

Design and Procedure. The selected 48 co-laughter stimuli were amplitude normalized and presented in random order using SuperLab 4.0 experiment software (www.superlab.com). For those study sites in which a language other than English was employed in conducting the experiment, the instructions were translated beforehand by the respective investigators, or by native-speaker translators recruited by them for this purpose (see Table 4 for participants' native language information). Customized versions of the experiment were then created for each of these study sites using the translated instructions and a run-only version of the software. For those study sites in which literacy was limited or absent, the experimenter read the instructions aloud to each participant in turn.

Table 4. Demographic characteristics of study sites (see SI for discussion)

Country *Ethnic group*	**Participants' native language**	**Language in which experiment was conducted**	**Typical participant's English fluency**	**Mass media exposure**	**~% mass media in English**	**Typical participant's education**	**Degree of gender segregation**	**Community or city scale (number of people)**	**Economic mode(s) of participants**
Austria	German	German	moderate	extensive	50	Some college	none	small towns (<5,000) and large cities (>500,000)	Industrial: low and highly skilled
Belgium	Dutch	Dutch	moderate	extensive	75	Some college	none	medium cities (150,000-500,000)	Industrial: low and highly skilled
Brazil	Portuguese	Portuguese	minimal	extensive	<25	8-12 years	none	small (10,000-150,000) and large cities	Small-scale trade industrial: low and highly skilled
China	Chinese	Chinese	moderate	daily	50	Some college	none	large cities	Industrial: low and highly skilled
Denmark	Danish	Danish	moderate	extensive	50	Some college	none	medium cities	Industrial: highly skilled
Egypt	Egyptian Arabic	Egyptian Arabic	moderate	extensive	50	College degree	moderate	large cities	Industrial: highly skilled
India	Kannada	English	moderate	extensive	50	Some college	minimal	medium cities	Industrial: highly skilled
Indonesia	Jakarta dialect of Indonesian	Formal Indonesian	moderate	extensive	75	8-12 years	minimal	large cities	Small-scale trade industrial: low and highly skilled
Italy	Italian	Italian	minimal	extensive	<25	Some college	none	large cities	Industrial: highly skilled
Japan	Japanese	Japanese	minimal	daily	<25	8-12 years	minimal	large cities	Industrial: low and highly skilled
Korea	Korean	Korean	moderate	extensive	25	Some college	none	large cities	Small-scale: trade industrial: low and highly skilled
Namibia *Himba*	Otjiherero	Otjiherero	none	rare	25	1-3 years	minimal	small bands	Small-scale: horticulture, pastoralism

Netherlands	Dutch	Dutch	moderate	extensive	100	Some college	minimal	small and medium cities	Industrial: highly skilled
New Guinea *Sursurunga*	Sursurunga	Neo-Melanesian	minimal	rare	75	4-7 years	minimal	small villages	Small-scale horticulture
New Zealand	English	English	primary language	extensive	100	College degree	none	large cities	Industrial: highly skilled
Peru (rural)	Spanish	Spanish	minimal	extensive	<25	8-12 years	none	small towns	Small-scale horticulture, agriculture, pastoral industrial: low skill
Peru (urban)	Spanish	Spanish	moderate	extensive	50	Some college	none	large cities	Industrial: highly skilled
Singapore	English	English	primary language	extensive	75	Some college	none	large cities	Industrial: low and highly skilled
Slovakia	Slovak	Slovak	minimal	extensive	<25	College degree	none	large towns (5,000-10,000) and small cities	Industrial: highly skilled
South Africa *Zulu*	isiZulu	isiZulu	minimal	occasional	<25	8-12 years	minimal	large villages (200-1,000) and small towns	Small-scale agriculture, pastoralism, trade industrial: low-skill
Spain	Spanish	Spanish	minimal	extensive	<25	Some college	none	small and medium cities	Industrial: low and highly skilled
Tanzania *Hadza*	Hadzane	Hadzane and Swahili	none	rare	<25	1-3 years	moderate	small bands	Hunter-gatherer
Turkey	Turkish	Turkish	moderate	extensive	50	8-12 years	minimal	large cities	Small-scale trade industrial: low and highly skilled
USA	English	English	primary language	extensive	100	Some college	none	large cities	Industrial: highly skilled

Prior to each experiment, participants were instructed that they would be listening to recordings of pairs of people laughing together in a conversation, and they would be asked questions about each recording. Participants received one practice trial and then completed the full experiment consisting of 48 trials. After each trial, listeners answered two questions. The first question was a two-alternative forced-choice asking them to identify the pair as either friends or strangers, while the second question asked listeners to judge how well the pair liked one another on a scale of 1 to 7, with 1 being *not at all*, and 7 being *very much*. The scale was presented visually, and, in study sites where the investigator judged participants' experience with numbers and/or scales to be low, participants used their finger to point to the appropriate part of the scale. All participants wore headphones. See SI for complete text of instructions and questions employed in the experiment.

ACKNOWLEDGEMENTS

We thank our participants around the globe, and Brian Kim and Andy Lin of the UCLA Statistical Consulting Group.

The data reported in this paper are archived at https://escholarship.org/uc/item/99j8r0gx

Author contributions: the first four authors are listed in order of the importance of their contributions. GB designed the hypothesis and methods, conducted the core analyses, and wrote the draft manuscript. DF envisioned the cross-cultural component, organized the cross-cultural research, and assisted in writing the manuscript. RF conducted the acoustic analyses and contributed corresponding draft text. EC managed the cross-cultural research. All remaining authors contributed data, and are listed in alphabetical order.

REFERENCES

1. Dale R, Fusaroli R, Duran N, Richardson D (2013) The self-organization of human interaction. *Psychol Learn Motiv* 59: 43-95
2. Provine R R (2000) *Laughter: A scientific investigation* Penguin Press, New York, NY.
3. Gervais M, Wilson D (2005) The evolution and functions of laughter and humor: A synthetic approach *Q Rev Biol* 80(4): 395–430.
4. Scott SK, Lavan N, Chen S, McGettigan C (2014) The social life of laughter. *Trends Cogn Sci* 18, 618-620.
5. Sauter DA, Eisner F, Ekman P, Scott SK (2010) Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proc Natl Acad Sci* 107(6): 2408–2412.
6. Smoski MJ, Bachorowski JA (2003) Antiphonal laughter between friends and strangers. *Cognition Emotion* 17(2): 327-340.
7. Vettin J, Todt D (2004) Laughter in conversation: Features of occurrence and acoustic structure *J Nonverbal Behav* 28(2): 93–115.
8. Luschei ES et al (2006) Patterns of laryngeal EMG and the activity of the respiratory system during spontaneous laughter. *J Neurophysiol* 96(1): 442–450.
9. Titze IR et al (2008) Laryngeal muscle activity in giggle: A damped oscillation model. *J Voice* 22(6): 644–648.
10. Bachorowski JA, Smoski MJ, Owren MJ (2001) The acoustic features of human laughter. *J Acoust Soc Am* 110(3): 1581–1597.
11. Provine RR (1993) Laughter punctuates speech: Linguistic, social and gender contexts of laughter. *Ethology* 95(4): 291–298.
12. Davila-Ross M, Owren MJ, Zimmermann E (2009) Reconstructing the evolution of laughter in great apes and humans. *Curr Biol* 19(13): 1106–1111.
13. van Hooff JA (1972) A comparative approach to the phylogeny of laughter and smiling. *Nonverbal communication,* ed Hinde RA (Cambridge University Press, Cambridge, England), pp 209–241.
14. Bachorowski JA, Owren MJ (2001) Not all laughs are alike: Voiced but not unvoiced laughter readily elicits positive affect. *Psychol Sci* 12(3): 252–257.
15. Kipper S, Todt D (2001) Variation of sound parameters affects the evaluation of human laughter. *Behaviour*, 138(9): 1161–1178.
16. Bryant GA, Aktipis CA (2014) The animal nature of spontaneous human laughter. *Evol Hum Behav* 35(4): 327-335.
17. McGettigan C et al (2015) Individual differences in laughter perception reveal roles for mentalizing and sensorimotor systems in the evaluation of emotional authenticity. *Cereb Cortex* bht227.
18. Szameitat DP et al (2010) It is not always tickling: Distinct cerebral responses during perception of different laughter types. *Neuroimage* 53(4): 1264–1271.
19. Jurgens U (2002) Neural pathways underlying vocal control. *Neurosci Biobehav R* 26(2): 235–238.
20. Ackermann H, Hage SR, Ziegler W (2014) Brain mechanisms of acoustic communication in humans and nonhuman primates: An evolutionary perspective. *Behav Brain Sci* 37(6): 529-546.

21. Ambady N, Bernieri FJ, Richeson JA (2000) Toward a histology of social behavior: Judgmental accuracy from thin slices of the behavioral stream. *Adv Exp Soc Psychol* 32: 201-271.
22. Henrich J, Heine SJ, Norenzayan A (2010) The weirdest people in the world? *Behav Brain Sci* 33(2-3): 61–83.
23. Bates D, Maechler M, Bolker B, Walker (2014) lme4: Linear mixed-effects models using Eigen and S4 R package version 11-7, http://CRANR-projectorg/package=lme4
24. R Core Team (2014) R: A language and environment for statistical computing R Foundation for Statistical Computing, Vienna, Austria URL http://wwwR-projectorg/
25. Zou H, Hastie T (2005) Regularization and variable selection via the elastic net. *J Roy Stat Soc B* 67(2): 301-320.
26. Williams CE, Stevens, KN (1972) Emotions and speech: Some acoustical correlates. *J Acoust Soc Am* 52(4B): 1238-1250.
27. Bryant GA (2010) Prosodic contrasts in ironic speech. *Discourse Processes* 47(7): 545-566.
28. Boersma P, Weenink D (2014) Praat: Doing phonetics by computer [Computer program] Version 54, retrieved 4 October 2014 from http://wwwpraatorg/
29. Fusaroli R, Konvalinka L, Wallot S (2014) Analyzing social interactions: Promises and challenges of cross recurrence quantification analysis. *Springer Proceedings in Mathematics & Statistics.*
30. Marwan N, Carmen Romano M, Thiel M, Kurths (2007) Recurrence plots for the analysis of complex systems *Phys Rep* 438(5): 237-329.
31. Tsanas A, Little MA, McSharry PE, Ramig LO (2011) Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson's disease symptom severity *J Roy Soc Interface*, 8(59): 842-855.
32. Takens F (1981) Detecting strange attractors in turbulence. *Dynamical systems and turbulence, Warwick 1980* (Springer, Berlin), pp 366-381.
33. Zou H, Hastie T (2005) Regularization and variable selection via the elastic net *J Roy Stat Soc B* 67, 301-320.

Supporting Information

Bryant et al.

Participants. 966 adults participated in the study (412 men and 554 women). At one study site (Tanzania/Hadza), listeners were not asked to provide a rating (Question 2) due to unfamiliarity with such tasks. In Papua New Guinea, four participants did not finish the experiment. An additional six participants in the Papua New Guinea sample answered "friends" on every question. Because the instructions specified that some of the pairs were friends and some were newly-acquainted strangers, these six participants were judged to have either misunderstood the instructions or failed to follow them, and were therefore removed prior to the analysis. One Himba participant similarly answered "friends" every time and was therefore also removed prior to the analysis.

Table S1. Sex and age breakdown by study site

Group	Region	N	Men	Women	Mean Age	Age SD	Age Range
South Africa/Zulu	Africa	100	50	50	29.3	9.2	19-53
Egypt	Africa	29	7	22	30.2	10.3	18-61
Namibia/Himba	Africa	19	10	9	36.6	20.8	16-78
Tanzania/Hadza	Africa	53	29	24	40.5	12.1	21-71
Singapore	Asia	43	13	30	20.8	1.5	19-24
Korea	Asia	76	21	55	22.3	1.8	20-27
Japan	Asia	28	14	14	19.9	1.0	19-23
India	Asia	38	20	18	24.2	3.5	21-39
China	Asia	21	8	13	NA	NA	NA
Indonesia	Asia	36	15	21	28.5	11.2	18-62
Denmark	Europe	33	17	16	22.5	2.1	19-26
Netherlands	Europe	40	24	16	22.2	2.2	18-27
Belgium	Europe	26	5	21	NA	NA	18-25
Slovakia	Europe	40	13	27	29.6	10.3	21-59
Turkey	Europe	30	16	14	23.9	8.8	19-68
Spain	Europe	34	19	15	27.0	9.7	18-57
Austria	Europe	37	14	23	27.2	7.5	19-49
Italy	Europe	45	24	21	32.3	14.5	19-62
USA	N. America	72	14	58	19.1	1.6	18-23
New Guinea/Sursurunga	Oceania	25	14	11	37.8	13.6	21-65
New Zealand	Oceania	44	23	21	31.9	10.2	19-71
Peru (urban)	S. America	30	9	21	20.8	2.3	18-27
Peru (rural)	S. America	31	12	19	31.1	11.2	18-58
Brazil	S. America	36	21	15	31.6	8.1	17-46
TOTALS	**6**	**966**	**412**	**554**	**27.7**	**7.9**	**18.8-47.9**

Table S2. Model comparisons for Question 1 (Friends or strangers?): 24 societies, 2X3 design

Model	Fixed factors	Random factors	Estimate	SE	z	Variance	SD	AIC
M1	(Intercept)		0.18187	0.01458	12.47			61082.7
	Condition1		0.56259	0.01936	29.06			
		Participant				0.03588	0.1894	
M2	(Intercept)		0.18364	0.03560	5.158			60901.1
	Condition1		0.56261	0.01935	29.07			
		Participant				0.00830	0.0918	
		Society				0.02565	0.1602	
M3	(Intercept)		0.15615	0.03704	4.216			60897.8
	Condition1		0.56261	0.01935	29.07			
	Sex		0.04852	0.02102	2.309			
		Participant				0.00783	0.0885	
		Society				0.02479	0.1574	
M4	(Intercept)		0.13788	0.03852	3.579			60896.8
	Condition1		0.60140	0.02953	20.366			
	Sex		0.08050	0.02791	2.884			
	Condition1 x Sex		-0.06808	0.03909	-1.742			
		Participant				0.00783	0.0885	
		Society				0.02480	0.1575	
M5	(Intercept)		0.06329	0.04203	1.506			60758.6
	Condition1		0.60332	0.02956	20.410			
	Sex		0.08077	0.02801	2.888			
	Condition1 x Sex		-0.06833	0.03910	-1.729			
	ConvType1		-0.03810	0.02532	-1.505			
	ConvType2		0.21068	0.02447	8.611			
		Participant				0.00817	0.0904	
		Society				0.02496	0.1580	
M6	(Intercept)		0.09387	0.04497	2.087			60763.1
	Condition1		0.56440	0.01939	29.111			
	Sex		0.02722	0.03955	0.688			
	ConvType1		-0.06075	0.03865	-1.572			
	ConvType2		0.19953	0.03731	5.347			
	ConvType1 x Sex		0.03972	0.05117	0.776			
	ConvType2 x Sex		0.01952	0.04943	0.395			
		Participant				0.00816	0.0903	
		Society				0.02494	0.1579	
M7	**(Intercept)**		0.53231	0.04585	11.61			**58172.7**
	Condition1		-0.34731	0.03817	-9.10			
	Sex		0.05142	0.02227	2.31			
	ConvType1		-0.15106	0.03600	-4.20			
	ConvType2		-0.77395	0.03429	-22.57			
	Condition1 x ConvType1		0.22228	0.05034	4.42			
	Condition1 x ConvType2		2.21152	0.05147	42.96			
		Participant				0.01440	0.120	
		Society				0.02789	0.167	

Table S3. Model comparisons for Question 2 (Ratings of liking): 23 societies, 2X3 design

Model	**Fixed factors**	**Random factors**	**Estimate**	**SE**	**t**	**Variance**	**SD**	**AIC**
M1	(Intercept)		3.81056	0.0228	167.47			168916.6
	Condition1		0.83192	0.0155	53.50			
		Participant				0.3628	0.602	
M2	(Intercept)		3.81251	0.0415	92.87			168894.2
	Condition1		0.83192	0.0155	53.50			
		Participant				0.33916	0.582	
		Society				0.02626	0.163	
M3	(Intercept)		3.78484	0.0475	79.65			168894.9
	Condition1		0.83192	0.0155	53.22			
	Sex		0.04984	0.0432	1.16			
		Participant				0.33861	0.582	
		Society				0.02624	0.162	
M4	(Intercept)		3.81087	0.0483	78.86			168888.1
	Condition1		0.77982	0.0235	33.20			
	Sex		0.00351	0.0459	0.08			
	Condition1 x Sex		0.09272	0.0313	2.96			
		Participant				0.33859	0.582	
		Society				0.02624	0.162	
M5	(Intercept)		3.36451	0.0500	67.32			166281.4
	Condition1		0.77982	0.0228	34.22			
	Sex		0.00361	0.0457	0.08			
	Condition1 x Sex		0.09540	0.0305	3.13			
	ConvType1		0.21500	0.0200	10.78			
	ConvType2		0.89914	0.0191	47.13			
		Participant				0.34188	0.585	
		Society				0.02624	0.162	
M6	(Intercept)		3.34490	0.0514	65.05			166292.5
	Condition1		0.83192	0.0151	55.15			
	Sex		0.03850	0.0506	0.76			
	ConvType1		0.20912	0.0302	6.94			
	ConvType2		0.88842	0.0288	30.82			
	ConvType1 x Sex		0.01045	0.0402	0.26			
	ConvType2 x Sex		0.01909	0.0385	0.50			
		Participant				0.34187	0.585	
		Society				0.02624	0.162	
M7	**(Intercept)**		3.38587	0.0510	66.49			**166042.4**
	Condition1		0.73713	0.0301	24.51			
	Sex		0.05000	0.0432	1.16			
	ConvType1		0.30152	0.0281	10.72			
	ConvType2		0.71617	0.0270	26.62			
	Condition1 x ConvType1		-0.17306	0.0400	-4.35			
	Condition1 x ConvType2		0.32595	0.0381	9.62			
		Participant				0.34217	0.585	
		Society				0.02624	0.162	

Fig. S1. Rating responses (Y-axis) to Question 2 (On a scale of 1 to 7: How well do the speakers like one another?) in each study site broken down by experimental condition (strangers clustered on the left, or friends clustered on the right), and dyad type (male-male, male-female, female-female). Tanzania/Hadza were not asked Question 2 and do not appear here. See Table S7 for means and standard deviation values.

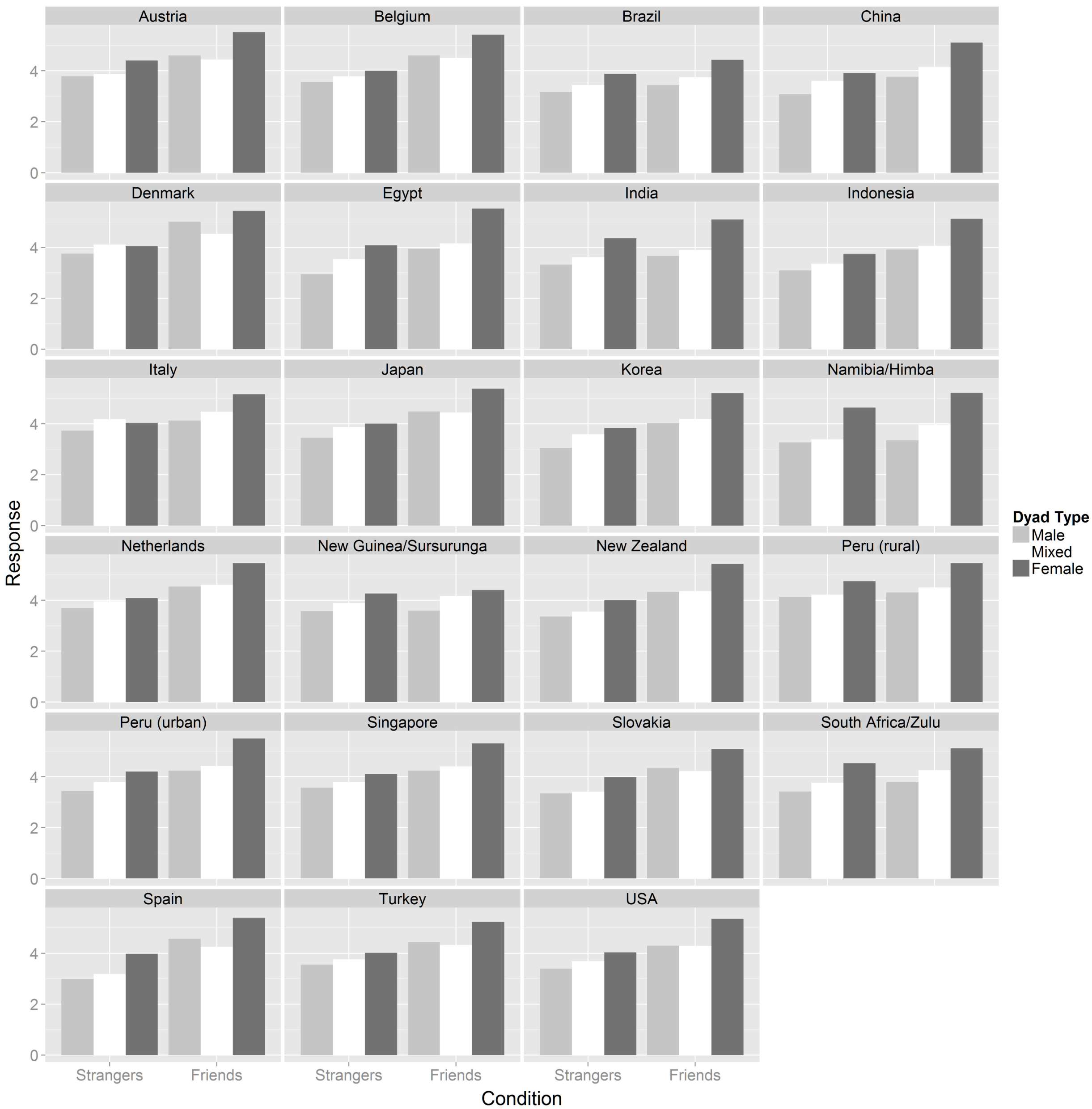
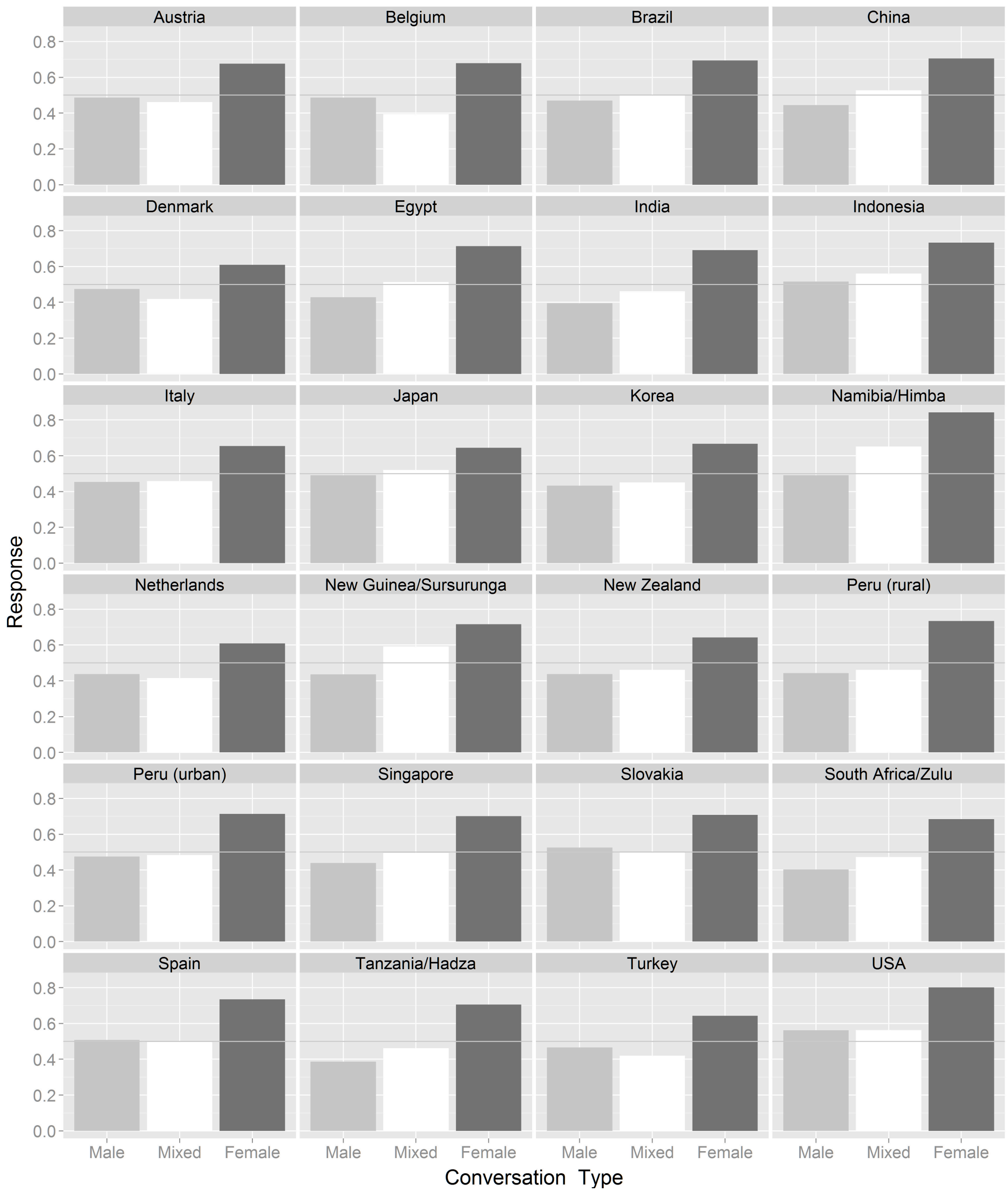


Fig. S2. Mean response rates (0 to 1) on Question 1 across all societies revealing universal bias to respond with answer of "friends" in female-female dyads, but not other dyad types. Y-axis reference line represents unbiased response rate (0.5).



Acoustic analysis of co-laughter segments

Measures

Having demonstrated that participants could accurately judge whether co-laughers are friends or strangers, we then measured a wide range of acoustic features of the laughter to identify which features would best explain the variance in participants' judgments. We examined the 48 co-laughter segments used in the experiment, for a total of 96 individual laughs. Two laughs were excluded (pair 3, participant 1, late laugh; pair 14, participant 2, late laugh) as they did not contain sufficient voicing (i.e., periodic) duration to allow us to automatically assess all of the acoustic features examined. We then analyzed the remaining 94 laughs, 47 of which were produced by friends and 47 of which were produced by strangers.

For each individual laugh within a given audio clip we measured the rate of intervoicing interval (rate of IVI) (16). We first calculated bout duration for each laugh from the onset of visible acoustic energy as viewed in a spectrogram (FFT method, window length: 0.005 s., time steps: 1000, frequency steps: 250, Gaussian window shape, dynamic range: 50 dB) to the offset of energy in the final call, or bout-final inspiratory element. Calls were counted based on audible and visible separated voiced energy. Mean call duration was calculated as total bout duration divided by call number. Mean intervoicing interval (IVI) was calculated as the summed lengths of all unvoiced intervals between calls (i.e., voiced call offset to voice call onset) divided by call number minus one. Unvoiced portions were determined by a lack of formant structure as viewed through a spectrogram with settings described above, and lack of periodicity with standard pitch range values. Finally, rate of IVI was calculated using the following formula:

$$\frac{\left(\frac{(\sum x_i)}{(c-1)}\right)}{\left(\frac{d}{c}\right)}$$

where x_i are the inter-voicing interval values, c is the total call number, and d is the bout duration of the series. This measure captures the averaged rate of unvoiced segments per call across a laugh bout. We also calculated the amount of overlap between co-laughs, that is, the duration for which both laughs can be heard at the same time (co-laugh overlap). Onsets and offsets of each laugh were automatically extracted by individuating the first and last data points in the signal with intensity >30 db, minimizing extraction of possible leakage from each laugh's counterpart. Overlap was calculated as the difference between the earliest offset and the latest onset in each co-laugh pair. Co-laughter between friends and strangers did not differ in overlap as indicated by a mixed model with Overlap as the dependent variable, Familiarity as the fixed factor, and Pair as a random intercept: $\beta = 0.08$, $SE = 0.14$, $t = 0.60$, $P = 0.55$. Nonetheless, in the interests of maximal rigor, we opted to include this variable in the model described below.

Using Praat (*28*), we extracted fundamental frequency (F_0) (frequency range = 70-400 Hz), and intensity during voiced intervals. F_0 values were converted to a logarithmic scale to approximate perceptual pitch, where after we produced sequences of voiced (presence of pitch) and unvoiced segments. Per each of these measures we calculated traditional descriptive statistics and temporal dynamics measures.

Descriptive statistics. We calculated a) the total and voiced duration of each laugh and the number of voiced and unvoiced segments, and b) standard deviations of pitch and intensity, as well as the mean and standard deviation of length of voiced and unvoiced segments.

Temporal dynamics measures. Traditional descriptive statistics do not capture other crucial aspects of time-series properties such as their regularity over time and the temporal-dependence between successive data points. These properties express the stability and complexity of voice production and have proven particularly useful to assess vocal behavior in a wide variety of contexts (for a review see [29]). To assess these temporal dynamics we employed two non-linear methods: a) Recurrence Quantification Analysis (RQA) of both voiced/unvoiced sequences and pitch (30); and b) Teager–Kaiser energy operator of pitch (31). RQA is a general non-linear time-series analysis tool that quantifies multiple aspects of the temporal stability of a time series, such as how repetitive, noisy, or stationary it is.

Relying on the time series in each laugh (e.g., a sequence of estimated pitch regularly sampled over time), RQA reconstructs the phase space of possible combinations of states and quantifies the structure of recurrence; that is, the number of instances in which the time series displays repeated dynamics, and the characteristics of these repetitions. In order to apply RQA, two steps are necessary: 1) reconstructing the phase space underlying the time series and 2) production of a recurrence plot. The phase space of a time series is an n-dimensional space in which all possible states of a system are represented, so that it is possible to portray the trajectories of the system's behavior, be it periodic (repeatedly crossing the same regions at regular intervals), random, or chaotic. In order to reconstruct the phase space, we applied the time-delay method (32) to each time series. After reconstructing the phase space, we constructed recurrence plots for each time series. Black dots on the plots represent every occasion at which a phase space trajectory goes through approximately the same region in the phase space. In mathematical terms, if we represent the trajectory of a system as

$$\{\vec{x_i}\}_{i=1}^{N}$$

the corresponding recurrence plot is based on the following recurrence matrix:

$$R_{i,j} = \begin{cases} 1: \vec{x_i} \approx \vec{x_j}, \\ 0: \vec{x_i} \not\approx \vec{x_j}, \end{cases} i, j = 1, \ldots, N$$

where N is the number of considered states of the system and $\vec{x_i} \approx \vec{x_j}$ indicates that the two states are equal up to an error (or distance) ε. Note that this ε is essential in the case of continuous variables (as in F_0) as systems often do not recur exactly, but only approximately revisit states. To statistically analyze differences in laughs, we performed Recurrence Quantification Analysis (RQA) on the recurrence plots. RQA provides several indices quantifying the structure and complexity of dynamical systems from recurrence plots (30). This makes it possible to statistically compare different dynamic systems (e.g., different dyads) in terms of their dynamics such as the stability, structure, and complexity in the behavior of the system. In particular we analyzed:

Amount of repetition: The percentage of values that recur (are repeated) in the time series independently of the lag (recurrence rate, RR).

$$RR(\varepsilon) = \frac{1}{N^2} \Sigma_{i,j=1}^{N} R_{i,j}(\varepsilon)$$

Stability of repetition: articulated in:

Average length of sequences repeated (L)

$$L = \frac{\sum_{l=lmin}^{N} l \, P(l)}{\sum_{l=lmin}^{N} P(l)}$$

Length of longest repeated sequence (LMAX)

$$LMAX = \max(\{l_i\}_{i=1}^{N_l})$$

For more details about these indices see (29).

The Teager–Kaiser energy operator (TKEO) has been widely employed to quantify jitter and shimmer; that is, perturbations in the regular cycles of pitch and intensity, respectively, which often characterize situations of stress and arousal, and are impacted by the ability to control the speech production system (31). TKEO is calculated as

$$\psi(x_n) = x_n^2 - x_{n+1} \cdot x_{n-1}$$

where the subscript n denotes the n^{th} entry of the vector x (in our case, the time series of pitch). We computed the mean, standard deviation and 5th, 25th, 75th and 95^{th} percentile values of TKEO.

Overall, this resulted in 34 features for each individual laugh.

Table S4. 34 extracted features in acoustic analysis.

Voiced / Unvoiced segments	**Pitch**	**Intensity**
Co-laugh overlap	Recurrence Rate (RR)	Recurrence Rate (RR)
Total duration of the laugh	Average length of recurrent sequence (L)	Average length of recurrent sequence (L)
Voiced duration	Maximum length of recurrent sequence (LMAX)	Maximum length of recurrent sequence (LMAX)
Unvoiced duration	Pitch SD	Intensity SD
Number of unvoiced segments	Average TKEO	Average TKEO
Average length of unvoiced segments	SD of TKEO	SD of TKEO
Average length of voiced segments	5th percentile of TKEO	5th percentile of TKEO
SD of length of unvoiced segments	25th percentile of TKEO	25th percentile of TKEO
SD of length of voiced segments	75th percentile of TKEO	75th percentile of TKEO
Mean intervoicing interval	95th percentile of TKEO	95th percentile of TKEO
Mean call duration		
Recurrence Rate (RR)		
Average length of recurrent sequence (L)		
Maximum length of recurrent sequence (LMAX)		

Besides individual features in each laugh, we were also interested in the relationship between laughs in a co-laughter dyad: do co-laughs between friends share features that co-laughs between strangers do not? Therefore, we prepared a second dataset by calculating the difference in values across all features for each pair of laughs. This yielded a dataset of 34 features for each of the 46 pairs of laughs for which we had acoustic features of both of the individual laughs making up the pair. These two datasets were used separately to assess which acoustic features listeners might employ when judging whether a co-laugh was between friends or strangers. We call this measure the *Friends Ratio* (FR), defined as the overall likelihood of each single laugh being part of a co-laugh segment produced between individuals identified by participants as being friends. In order to examine cross-cultural reliability, we then employed the overall model to predict within-cultures FR and assessed the amount of variance explained through Adjusted R^2. All acoustic features were linearly transformed on a scale from 0 to 1 for better performance in the feature selection process.

2.2. Analysis and machine learning process

Feature selection. The previously described process produces a large set of features, exemplifying what is commonly termed the curse of dimensionality. In other words, the presence of a large number of features makes the statistical models both difficult to interpret and at risk of overfitting, producing results that are not generalizable. To address this, we used a common algorithm to select a parsimonious subset of features: the Elastic Net extension of the LASSO (33), which could in principle reduce overall accuracy, but increases the interpretability and generalizability of the results, that is, the ability to accurately describe new laughs with characteristics similar to the laughs in the current study.

Statistical models. In order to assess the overall model relying on the selected features, we used a 5-fold cross-validated multiple regression model to reconstruct the participants' likelihood of judging a given dyad of co-laughers to be friends (FR). The dataset was divided into 5 subsets (or folds) each containing a non-overlapping fifth of the pairs of co-laughers. A combination of 4 folds was used for feature selection and model fitting. The model was then assessed on the remaining fold. This procedure was repeated for all four possible combinations of folds. In this way the accuracy of the model was assessed only on data on which it had not been trained. We repeated the cross-validation process a total of 100 times, randomly permuting the data before splitting into training and testing subsets to ensure stability of the results across different random splits in 5 folds.

Acoustic analysis results

The statistical model employing acoustic features of individual laughs was able to statistically predict FR to a high degree: $R^2 = 0.43$ (CI: 0.29 0.57), *Adjusted* $R^2 = 0.42$ (CI: 0.28 0.56), $p = 0.0001$. Features selected were: 1) Mean call duration; 2) Pitch average TKEO (mean jitter); 3) Pitch SD TKEO (SD jitter); 4) Intensity 5th percentile TKEO (5th PCTL shimmer). See Table 2.

In summary, the findings suggest that individual laughs that had 1) shorter average call duration, 2) less regular pitch cycles, 3) less variation in pitch cycle regularity, and 4) less regular intensity cycles were more likely to be rated as having been produced by friends. The model remains quite consistent across cultures as it explains a significant portion of the variance in the FR within each culture. See Table S5 for R^2 and Adjusted R^2 values in each culture. Figure 3 displays a scatterplot showing the correlation between participants' friend response across all cultures and predicted values using acoustic features selected by the statistical model.

Table S5. Acoustic feature selection model performance predicting friend response across 24 cultures.

Group	Region	R^2	Adjusted R^2
South Africa/Zulu	Africa	0.37	0.37
Egypt	Africa	0.34	0.33
Namibia/Himba	Africa	0.05	0.04
Tanzania/Hadza	Africa	0.30	0.29
Singapore	Asia	0.38	0.37
Korea	Asia	0.36	0.36
Japan	Asia	0.16	0.15
India	Asia	0.32	0.31
China	Asia	0.32	0.31
Indonesia	Asia	0.29	0.28
Denmark	Europe	0.20	0.19
Netherlands	Europe	0.19	0.18
Belgium	Europe	0.36	0.35
Slovakia	Europe	0.38	0.38
Turkey	Europe	0.21	0.20
Spain	Europe	0.41	0.41
Austria	Europe	0.38	0.37
Italy	Europe	0.35	0.34
USA	N. America	0.32	0.32
New Guinea/Sursurunga	Oceania	0.21	0.20
New Zealand	Oceania	0.34	0.33
Peru (urban)	S. America	0.38	0.37
Peru (rural)	S. America	0.50	0.49
Brazil	S. America	0.33	0.32

Responses provided by Himba (Namibian) participants constitute an exception to the model's ability to successfully explain a large proportion of the variance in FR across cultures, as the model explained only 5% of the variance in FR responses in this subsample. A separate model was trained on the Himba sample exclusively that explained slightly increased variance in their FR responses: Adjusted $R^2 = 0.08$ (CI: 0.04 0.14), $p = 0.04$. Features selected were: 1) lower mean call duration, and 2) less variability in pitch cycle regularity. See Table S6.

Table S6. Sample coefficients from one run of the 5-fold cross-validated model on Himba sample.

Predictor	Beta (SE) Fold1	Beta (SE) Fold2	Beta (SE) Fold3	Beta (SE) Fold4	Beta (SE) Fold5
Intercept	0.890 (0.051)	0.944 (0.046)	0.873 (0.036)	0.885 (0.038)	0.924 (0.041)
Mean call duration	-0.336 (0.108)	-0.382 (0.103)	-0.381 (0.088)	-0.218 (0.093)	-0.570 (0.106)
Pitch TKEO SD	-0.391 (0.110)	-0.601 (0.119)	-0.298 (0.098)	-0.298 (0.103)	-0.242 (0.100)

The Himba participants exhibited a stronger bias to judge co-laughers to be friends (69%) than is true in any other culture. B. Scelza, the co-author who collected these data, observed that some participants were occasionally confused by the question of "friends versus strangers," as this dichotomy does not allow for the identification of pairs of laughers who know each other but are not friends. However, closer analysis of the relationship between Himba participants' responses to Question 1 (friends versus strangers) and their responses to Question 2 (how well the pair liked one another) did not reveal an unusual pattern, suggesting that Himba participants interpreted Question 1 in a similar manner to participants in the other cultures studied.

The model employing the difference in acoustic features between co-laughers' laughs was not able to statistically predict FR to any significant degree. One limitation in this analysis for assessing coordination between co-laughers is the circumscribed overall length of the co-laughs (approximately 1 s.). An evolved signal of close affiliation would likely reveal the capacity for coordinated production, and this may be difficult to assess in short bouts of co-laughter isolated from the normal discursive context wherein bouts occur repeatedly. Future work should therefore investigate longer bouts of co-laughter to examine whether affiliated speakers produce coordinated laughter in a way that strangers do not, thus constituting more solid evidence that a signaling system is in place. At present, the results confirm, however, an available cue of affiliation that can be employed with quite limited information to make accurate judgments about the relationship between pairs of friends in dialogue.

Cultural and linguistic demographic dimensions

In principle, the extent to which third-party listeners can determine the relationship between co-laughers could be a function of many aspects of the listeners. For example, if the language spoken substantially influences the form of laughter, then, given that the stimuli were all generated by Californian speakers of American English, we might expect that listeners' familiarity with English would influence the accuracy of their judgments in this regard. Familiarity plausibly derives not only from the language spoken by the listener, but also from exposure to mass media in which English is used. More broadly, cultural similarity independent of familiarity with English could play a role. Likewise, a wide variety of studies in psychology reveal that individuals who are highly educated resemble one another across cultures (22), hence listeners' level of education might plausibly contribute to their ability to accurately discern the relationships obtaining between the undergraduate students whose laughter constituted the stimuli. Additionally, because participants were asked to make assessments of stimuli provided by dyads of each sex, and by mixed-sex dyads, the degree of gender segregation characteristic of a society could conceivably influence participants' accuracy. If experience plays a role in the ability to judge the nature of relationships on the basis of laughter, then individuals from highly gender-segregated societies would plausibly have less experience with mixed-sex dyads and dyads of the opposite sex, and thus might be expected to be less accurate. Similarly, if experience is critical, then participants from societies organized on the basis of small groups would necessarily have less experience with the range of idiosyncratic laughter styles possible than would participants from large cities, and thus could be expected to be less accurate. The manner in which an experiment is implemented could also have an effect on participants' responses, including whether instructions were read aloud to the participant by the researcher, or read visually by the participant on a computer screen. With these considerations in mind, the investigators responsible for each study site in this project estimated the features of their participants relevant to the above considerations; said data were not collected directly from the participants themselves, in part because, for some of the samples employed, questions concerning these matters could have damaged investigator rapport with the participants and/or disrupted the research process. Table 4 presents a summary of this information. Importantly, as noted in the text, despite the substantial multidimensional variation across study sites summarized in Table 4, neither the identity of the participants nor the location of the study sites accounted for substantial variance in the key dependent measures, thus underscoring the universality of the cues of affiliation presented by co-laughter.

Table S7. Means and standard deviations of judgments of hit rate (Question 1).

	Strangers						Friends					
	Male		Mixed		Female		Male		Mixed		Female	
Country	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Austria	0.62	0.49	0.59	0.49	0.50	0.50	0.59	0.49	0.52	0.50	0.85	0.36
Belgium	0.69	0.46	0.72	0.45	0.52	0.50	0.67	0.47	0.50	0.50	0.87	0.33
Brazil	0.56	0.50	0.55	0.50	0.38	0.49	0.50	0.50	0.55	0.50	0.77	0.42
China	0.63	0.49	0.60	0.49	0.47	0.50	0.52	0.50	0.65	0.48	0.88	0.33
Denmark	0.70	0.46	0.66	0.47	0.58	0.49	0.65	0.48	0.50	0.50	0.80	0.40
Egypt	0.70	0.46	0.56	0.50	0.44	0.50	0.56	0.50	0.59	0.49	0.87	0.34
India	0.68	0.47	0.61	0.49	0.41	0.49	0.47	0.50	0.53	0.50	0.79	0.40
Indonesia	0.61	0.49	0.52	0.50	0.43	0.50	0.64	0.48	0.65	0.48	0.89	0.31
Italy	0.59	0.49	0.64	0.48	0.51	0.50	0.50	0.50	0.55	0.50	0.82	0.39
Japan	0.62	0.49	0.55	0.50	0.50	0.50	0.60	0.49	0.59	0.49	0.79	0.41
Korea	0.70	0.46	0.70	0.46	0.54	0.50	0.56	0.50	0.60	0.49	0.87	0.33
Namibia/Himba	0.49	0.50	0.41	0.49	0.19	0.40	0.47	0.50	0.72	0.45	0.88	0.33
Netherlands	0.67	0.47	0.69	0.46	0.62	0.49	0.54	0.50	0.52	0.50	0.84	0.37
PG/Sursurunga	0.61	0.49	0.46	0.50	0.32	0.47	0.49	0.50	0.65	0.48	0.76	0.43
New Zealand	0.69	0.47	0.65	0.48	0.55	0.50	0.56	0.50	0.57	0.50	0.84	0.37
Peru (rural)	0.58	0.49	0.62	0.49	0.36	0.48	0.47	0.50	0.54	0.50	0.83	0.37
Peru (urban)	0.67	0.47	0.63	0.48	0.44	0.50	0.62	0.49	0.60	0.49	0.87	0.34
Singapore	0.67	0.47	0.59	0.49	0.50	0.50	0.55	0.50	0.58	0.49	0.90	0.30
Slovakia	0.59	0.49	0.58	0.49	0.44	0.50	0.64	0.48	0.58	0.49	0.86	0.35
South Africa/Zulu	0.63	0.48	0.58	0.49	0.40	0.49	0.43	0.50	0.52	0.50	0.77	0.42
Spain	0.70	0.46	0.62	0.49	0.39	0.49	0.71	0.45	0.62	0.49	0.86	0.35
Tanzania/Hadza	0.63	0.48	0.57	0.50	0.36	0.48	0.40	0.49	0.49	0.50	0.77	0.42

Turkey	0.65	0.48	0.62	0.49	0.52	0.50	0.58	0.49	0.46	0.50	0.80	0.40
USA	0.55	0.50	0.55	0.50	0.34	0.47	0.68	0.47	0.67	0.47	0.94	0.23

Table S8. Means and standard deviations of ratings of liking between co-laughers (Question 2).

	Strangers						Friends					
	Male		Mixed		Female		Male		Mixed		Female	
Country	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Austria	3.78	1.49	3.87	1.48	4.41	1.58	4.60	1.56	4.44	1.51	5.52	1.22
Belgium	3.55	1.53	3.78	1.53	4.00	1.47	4.61	1.55	4.51	1.41	5.42	1.30
Brazil	3.17	2.06	3.45	2.13	3.88	2.16	3.44	2.13	3.74	2.12	4.43	2.11
China	3.08	1.55	3.60	1.62	3.91	1.43	3.76	1.63	4.15	1.51	5.11	1.27
Denmark	3.75	1.48	4.10	1.47	4.05	1.45	5.01	1.44	4.52	1.46	5.43	1.23
Egypt	2.95	1.74	3.53	1.85	4.08	1.94	3.94	2.15	4.15	1.88	5.52	1.65
India	3.32	1.85	3.61	1.76	4.35	1.83	3.67	1.83	3.89	1.88	5.09	1.68
Indonesia	3.09	1.58	3.36	1.67	3.74	1.64	3.91	1.75	4.06	1.77	5.12	1.48
Italy	3.73	1.72	4.18	1.63	4.04	1.71	4.12	1.73	4.47	1.62	5.15	1.49
Japan	3.45	1.43	3.88	1.37	4.01	1.43	4.48	1.49	4.44	1.40	5.38	1.18
Korea	3.05	1.49	3.59	1.46	3.83	1.52	4.03	1.53	4.19	1.42	5.20	1.30
Namibia/Himba	3.27	2.25	3.39	2.15	4.64	2.02	3.35	2.29	3.97	2.07	5.22	1.77
Netherlands	3.70	1.45	3.96	1.45	4.08	1.48	4.54	1.47	4.60	1.48	5.45	1.21
PG/Sursurunga	3.57	1.92	3.90	1.95	4.26	1.97	3.59	2.03	4.17	1.97	4.40	1.98
New Zealand	3.36	1.67	3.55	1.63	4.00	1.68	4.33	1.73	4.36	1.66	5.43	1.39
Peru (rural)	4.13	1.70	4.22	1.56	4.75	1.61	4.31	1.64	4.50	1.65	5.45	1.30
Peru (urban)	3.45	1.62	3.80	1.64	4.20	1.63	4.24	1.60	4.42	1.49	5.50	1.43
Singapore	3.57	1.56	3.80	1.58	4.11	1.57	4.24	1.53	4.40	1.45	5.31	1.27

Slovakia	3.34	1.82	3.41	1.72	3.99	1.80	4.34	1.94	4.23	1.80	5.09	1.60
South Africa/Zulu	3.42	1.99	3.77	2.04	4.53	1.91	3.78	1.98	4.26	2.02	5.12	1.76
Spain	2.99	1.64	3.18	1.62	3.98	1.82	4.57	1.81	4.25	1.78	5.39	1.38
Turkey	3.54	1.75	3.76	1.66	4.01	1.62	4.43	1.77	4.33	1.73	5.24	1.53
USA	3.39	1.60	3.69	1.57	4.04	1.66	4.30	1.72	4.28	1.62	5.34	1.33

Complete text of English instructions for participants

The following text was used in the computerized experiment for English speakers, or as the basis for translation. The numbering below denotes separate screen presentation in the SuperLab experiment.

1) Welcome to the laughter and friends study. In this experiment we will have you listen to recordings of people in pairs laughing together in a conversation, and then answer two questions about each recording.

2) Some of the pairs of people were friends at the time of the recording, and others were complete strangers who were meeting for the first time. We will ask whether you think the people laughing together were friends or were strangers, and then we will ask you how well you think the people liked each other.

3) The recordings of people laughing are very short, and do not include any conversation. Before we begin with the actual study, you will be able to practice with one recording so that you will be familiar with the procedure.

4) Press the space bar when you are ready to begin the practice session.

5) When you are ready, press the space bar to hear the practice recording.

6) Do you think these people laughing were friends at the time of the conversation, or were they strangers who had just met for the first time? Press 1 if you think that the people in the recording were friends and 0 if you think that the people in the recording were strangers.

7) How much do you think these people liked each other? Provide your estimate using this scale, where 1 means "Not at all," 4 means "Somewhat," and 7 means "Very much."

8) When you are ready, press the space bar to hear the recording. If you have any questions, please ask the experimenter. If you do not have any questions, press the enter key and the experiment will begin.

9) You have now listened to all of the recordings. Thank you for your participation. Please tell the experimenter that you are finished.

Audio samples

Six audio (.wav format) examples of laughter types, and **s**ix sample stimulus files are provided online (12 total). All laughter recorded by GAB except S2.

S1: Normal laugh with periodic structure and egressive airflow (human_laugh.wav)
S2: Chimpanzee "laugh" (*Pan troglodytes*) with noisy, aperiodic structure and alternating airflow (chimp_laugh.wav) (Courtesy of Robert Provine).
S3: Voiced laugh (voiced_laugh.wav)
S4: Unvoiced laugh (unvoiced_laugh.wav)
S5: Spontaneous laughter (spontaneous_laugh.wav) (not from stimulus set of current study)
S6: Volitional laughter (volitional_laugh.wav) (not from stimulus set of current study)
S7: Friends, male-male pair (friends_mm.wav)
S8: Friends, male-female pair (friends_mf.wav)
S9: Friends, female-female pair (friends_ff.wav)
S10: Strangers, male-male pair (strangers_mm.wav)
S11: Strangers, male-female pair (strangers_mf.wav)
S12: Strangers, female-female pair (strangers_ff.wav)

SI Appendix

Detecting affiliation in co-laughter across 24 societies

Gregory A. Bryant, Daniel M. T. Fessler, Riccardo Fusaroli, Edward Clint, Lene Aarøe, Coren L. Apicella, Michael Bang Petersen, Shaneikiah T. Bickham, Alexander H. Bolyanatz, Brenda Chavez, Delphine De Smet, Cinthya Díaz, Jana Fančovičová, Michal Fux, Paulina Giraldo-Perez, Anning Hu, Shanmukh V. Kamble, Tatsuya Kameda, Norman P. Li, Francesca R. Luberti, Pavol Prokop, Katinka Quintelier, Brooke Scelza, Hyun Jung Shin, Montserrat Soler, Stefan Stieger, Wataru Toyokawa, Ellis A. van den Hende, Hugo Viciana-Asensio, Saliha Elif Yildizhan, Jose C. Yong, Tessa Yuditha, Yi Zhou

Table S1. Model comparisons for Question 1 (Friends or strangers?): 24 societies, 2X3 design

Model	Fixed factors	Random factors	Estimate	SE	z	Variance	SD	AIC
M1	(Intercept)		0.18187	0.01458	12.47			61082.7
	Condition1		0.56259	0.01936	29.06			
		Participant				0.03588	0.1894	
M2	(Intercept)		0.18364	0.03560	5.158			60901.1
	Condition1		0.56261	0.01935	29.07			
		Participant				0.00830	0.0918	
		Society				0.02565	0.1602	
M3	(Intercept)		0.15615	0.03704	4.216			60897.8
	Condition1		0.56261	0.01935	29.07			
	Sex		0.04852	0.02102	2.309			
		Participant				0.00783	0.0885	
		Society				0.02479	0.1574	
M4	(Intercept)		0.13788	0.03852	3.579			60896.8
	Condition1		0.60140	0.02953	20.366			
	Sex		0.08050	0.02791	2.884			
	Condition1 x Sex		-0.06808	0.03909	-1.742			
		Participant				0.00783	0.0885	
		Society				0.02480	0.1575	
M5	(Intercept)		0.06329	0.04203	1.506			60758.6
	Condition1		0.60332	0.02956	20.410			
	Sex		0.08077	0.02801	2.888			
	Condition1 x Sex		-0.06833	0.03910	-1.729			
	ConvType1		-0.03810	0.02532	-1.505			
	ConvType2		0.21068	0.02447	8.611			
		Participant				0.00817	0.0904	
		Society				0.02496	0.1580	
M6	(Intercept)		0.09387	0.04497	2.087			60763.1
	Condition1		0.56440	0.01939	29.111			
	Sex		0.02722	0.03955	0.688			
	ConvType1		-0.06075	0.03865	-1.572			
	ConvType2		0.19953	0.03731	5.347			
	ConvType1 x Sex		0.03972	0.05117	0.776			
	ConvType2 x Sex		0.01952	0.04943	0.395			
		Participant				0.00816	0.0903	
		Society				0.02494	0.1579	
M7	**(Intercept)**		0.53231	0.04585	11.61			**58172.7**
	Condition1		-0.34731	0.03817	-9.10			
	Sex		0.05142	0.02227	2.31			
	ConvType1		-0.15106	0.03600	-4.20			
	ConvType2		-0.77395	0.03429	-22.57			
	Condition1 x ConvType1		0.22228	0.05034	4.42			
	Condition1 x ConvType2		2.21152	0.05147	42.96			
		Participant				0.01440	0.120	
		Society				0.02789	0.167	

Table S2. Model comparisons for Question 2 (Ratings of liking): 23 societies, 2X3 design

Model	Fixed factors	Random factors	Estimate	SE	t	Variance	SD	AIC
M1	(Intercept)		3.81056	0.0228	167.47			168916.6
	Condition1		0.83192	0.0155	53.50			
		Participant				0.3628	0.602	
M2	(Intercept)		3.81251	0.0415	92.87			168894.2
	Condition1		0.83192	0.0155	53.50			
		Participant				0.33916	0.582	
		Society				0.02626	0.163	
M3	(Intercept)		3.78484	0.0475	79.65			168894.9
	Condition1		0.83192	0.0155	53.22			
	Sex		0.04984	0.0432	1.16			
		Participant				0.33861	0.582	
		Society				0.02624	0.162	
M4	(Intercept)		3.81087	0.0483	78.86			168888.1
	Condition1		0.77982	0.0235	33.20			
	Sex		0.00351	0.0459	0.08			
	Condition1 x Sex		0.09272	0.0313	2.96			
		Participant				0.33859	0.582	
		Society				0.02624	0.162	
M5	(Intercept)		3.36451	0.0500	67.32			166281.4
	Condition1		0.77982	0.0228	34.22			
	Sex		0.00361	0.0457	0.08			
	Condition1 x Sex		0.09540	0.0305	3.13			
	ConvType1		0.21500	0.0200	10.78			
	ConvType2		0.89914	0.0191	47.13			
		Participant				0.34188	0.585	
		Society				0.02624	0.162	
M6	(Intercept)		3.34490	0.0514	65.05			166292.5
	Condition1		0.83192	0.0151	55.15			
	Sex		0.03850	0.0506	0.76			
	ConvType1		0.20912	0.0302	6.94			
	ConvType2		0.88842	0.0288	30.82			
	ConvType1 x Sex		0.01045	0.0402	0.26			
	ConvType2 x Sex		0.01909	0.0385	0.50			
		Participant				0.34187	0.585	
		Society				0.02624	0.162	
M7	**(Intercept)**		3.38587	0.0510	66.49			**166042.4**
	Condition1		0.73713	0.0301	24.51			
	Sex		0.05000	0.0432	1.16			
	ConvType1		0.30152	0.0281	10.72			
	ConvType2		0.71617	0.0270	26.62			
	Condition1 x ConvType1		-0.17306	0.0400	-4.35			
	Condition1 x ConvType2		0.32595	0.0381	9.62			
		Participant				0.34217	0.585	
		Society				0.02624	0.162	

Table S3. Means and standard deviations of judgments of hit rate (Question 1: “Do you think these people laughing were friends or strangers?”).

	Strangers						Friends					
	Male		Mixed		Female		Male		Mixed		Female	
Country	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Austria	0.62	0.49	0.59	0.49	0.50	0.50	0.59	0.49	0.52	0.50	0.85	0.36
Belgium	0.69	0.46	0.72	0.45	0.52	0.50	0.67	0.47	0.50	0.50	0.87	0.33
Brazil	0.56	0.50	0.55	0.50	0.38	0.49	0.50	0.50	0.55	0.50	0.77	0.42
China	0.63	0.49	0.60	0.49	0.47	0.50	0.52	0.50	0.65	0.48	0.88	0.33
Denmark	0.70	0.46	0.66	0.47	0.58	0.49	0.65	0.48	0.50	0.50	0.80	0.40
Egypt	0.70	0.46	0.56	0.50	0.44	0.50	0.56	0.50	0.59	0.49	0.87	0.34
India	0.68	0.47	0.61	0.49	0.41	0.49	0.47	0.50	0.53	0.50	0.79	0.40
Indonesia	0.61	0.49	0.52	0.50	0.43	0.50	0.64	0.48	0.65	0.48	0.89	0.31
Italy	0.59	0.49	0.64	0.48	0.51	0.50	0.50	0.50	0.55	0.50	0.82	0.39
Japan	0.62	0.49	0.55	0.50	0.50	0.50	0.60	0.49	0.59	0.49	0.79	0.41
Korea	0.70	0.46	0.70	0.46	0.54	0.50	0.56	0.50	0.60	0.49	0.87	0.33
Namibia/Himba	0.49	0.50	0.41	0.49	0.19	0.40	0.47	0.50	0.72	0.45	0.88	0.33
Netherlands	0.67	0.47	0.69	0.46	0.62	0.49	0.54	0.50	0.52	0.50	0.84	0.37
PG/Sursurunga	0.61	0.49	0.46	0.50	0.32	0.47	0.49	0.50	0.65	0.48	0.76	0.43
New Zealand	0.69	0.47	0.65	0.48	0.55	0.50	0.56	0.50	0.57	0.50	0.84	0.37
Peru (rural)	0.58	0.49	0.62	0.49	0.36	0.48	0.47	0.50	0.54	0.50	0.83	0.37
Peru (urban)	0.67	0.47	0.63	0.48	0.44	0.50	0.62	0.49	0.60	0.49	0.87	0.34
Singapore	0.67	0.47	0.59	0.49	0.50	0.50	0.55	0.50	0.58	0.49	0.90	0.30
Slovakia	0.59	0.49	0.58	0.49	0.44	0.50	0.64	0.48	0.58	0.49	0.86	0.35
South Africa/Zulu	0.63	0.48	0.58	0.49	0.40	0.49	0.43	0.50	0.52	0.50	0.77	0.42
Spain	0.70	0.46	0.62	0.49	0.39	0.49	0.71	0.45	0.62	0.49	0.86	0.35
Tanzania/Hadza	0.63	0.48	0.57	0.50	0.36	0.48	0.40	0.49	0.49	0.50	0.77	0.42
Turkey	0.65	0.48	0.62	0.49	0.52	0.50	0.58	0.49	0.46	0.50	0.80	0.40
USA	0.55	0.50	0.55	0.50	0.34	0.47	0.68	0.47	0.67	0.47	0.94	0.23

Table S4. Means and standard deviations of ratings of liking between co-laughers (Question 2: "How much do these people like each other?").

	Strangers						Friends					
	Male		Mixed		Female		Male		Mixed		Female	
Country	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Austria	3.78	1.49	3.87	1.48	4.41	1.58	4.60	1.56	4.44	1.51	5.52	1.22
Belgium	3.55	1.53	3.78	1.53	4.00	1.47	4.61	1.55	4.51	1.41	5.42	1.30
Brazil	3.17	2.06	3.45	2.13	3.88	2.16	3.44	2.13	3.74	2.12	4.43	2.11
China	3.08	1.55	3.60	1.62	3.91	1.43	3.76	1.63	4.15	1.51	5.11	1.27
Denmark	3.75	1.48	4.10	1.47	4.05	1.45	5.01	1.44	4.52	1.46	5.43	1.23
Egypt	2.95	1.74	3.53	1.85	4.08	1.94	3.94	2.15	4.15	1.88	5.52	1.65
India	3.32	1.85	3.61	1.76	4.35	1.83	3.67	1.83	3.89	1.88	5.09	1.68
Indonesia	3.09	1.58	3.36	1.67	3.74	1.64	3.91	1.75	4.06	1.77	5.12	1.48
Italy	3.73	1.72	4.18	1.63	4.04	1.71	4.12	1.73	4.47	1.62	5.15	1.49
Japan	3.45	1.43	3.88	1.37	4.01	1.43	4.48	1.49	4.44	1.40	5.38	1.18
Korea	3.05	1.49	3.59	1.46	3.83	1.52	4.03	1.53	4.19	1.42	5.20	1.30
Namibia/Himba	3.27	2.25	3.39	2.15	4.64	2.02	3.35	2.29	3.97	2.07	5.22	1.77
Netherlands	3.70	1.45	3.96	1.45	4.08	1.48	4.54	1.47	4.60	1.48	5.45	1.21
PG/Sursurunga	3.57	1.92	3.90	1.95	4.26	1.97	3.59	2.03	4.17	1.97	4.40	1.98
New Zealand	3.36	1.67	3.55	1.63	4.00	1.68	4.33	1.73	4.36	1.66	5.43	1.39
Peru (rural)	4.13	1.70	4.22	1.56	4.75	1.61	4.31	1.64	4.50	1.65	5.45	1.30
Peru (urban)	3.45	1.62	3.80	1.64	4.20	1.63	4.24	1.60	4.42	1.49	5.50	1.43
Singapore	3.57	1.56	3.80	1.58	4.11	1.57	4.24	1.53	4.40	1.45	5.31	1.27
Slovakia	3.34	1.82	3.41	1.72	3.99	1.80	4.34	1.94	4.23	1.80	5.09	1.60
South Africa/Zulu	3.42	1.99	3.77	2.04	4.53	1.91	3.78	1.98	4.26	2.02	5.12	1.76
Spain	2.99	1.64	3.18	1.62	3.98	1.82	4.57	1.81	4.25	1.78	5.39	1.38
Turkey	3.54	1.75	3.76	1.66	4.01	1.62	4.43	1.77	4.33	1.73	5.24	1.53
USA	3.39	1.60	3.69	1.57	4.04	1.66	4.30	1.72	4.28	1.62	5.34	1.33

Participants

966 adults participated in the study (412 men and 554 women) (Tables S5-S6). At one study site (Tanzania/Hadza), listeners were not asked to provide a rating (Question 2) due to unfamiliarity with such tasks. In Papua New Guinea, four participants did not finish the experiment. An additional six participants in the Papua New Guinea sample answered “friends” on every question. Because the instructions specified that some of the pairs were friends and some were newly-acquainted strangers, these six participants were judged to have either misunderstood the instructions or failed to follow them, and were therefore removed prior to the analysis. One Himba participant similarly answered “friends” every time and was therefore also removed prior to the analysis.

Table S5. Sex and age breakdown by study site

Group	Region	N	Men	Women	Mean Age	Age SD	Age Range
South Africa/Zulu	Africa	100	50	50	29.3	9.2	19-53
Egypt	Africa	29	7	22	30.2	10.3	18-61
Namibia/Himba	Africa	19	10	9	36.6	20.8	16-78
Tanzania/Hadza	Africa	53	29	24	40.5	12.1	21-71
Singapore	Asia	43	13	30	20.8	1.5	19-24
Korea	Asia	76	21	55	22.3	1.8	20-27
Japan	Asia	28	14	14	19.9	1.0	19-23
India	Asia	38	20	18	24.2	3.5	21-39
China	Asia	21	8	13	NA	NA	NA
Indonesia	Asia	36	15	21	28.5	11.2	18-62
Denmark	Europe	33	17	16	22.5	2.1	19-26
Netherlands	Europe	40	24	16	22.2	2.2	18-27
Belgium	Europe	26	5	21	NA	NA	18-25
Slovakia	Europe	40	13	27	29.6	10.3	21-59
Turkey	Europe	30	16	14	23.9	8.8	19-68
Spain	Europe	34	19	15	27.0	9.7	18-57
Austria	Europe	37	14	23	27.2	7.5	19-49
Italy	Europe	45	24	21	32.3	14.5	19-62
USA	N. America	72	14	58	19.1	1.6	18-23
New Guinea/Sursurunga	Oceania	25	14	11	37.8	13.6	21-65
New Zealand	Oceania	44	23	21	31.9	10.2	19-71
Peru (urban)	S. America	30	9	21	20.8	2.3	18-27
Peru (rural)	S. America	31	12	19	31.1	11.2	18-58
Brazil	S. America	36	21	15	31.6	8.1	17-46
TOTALS	**6**	**966**	**412**	**554**	**27.7**	**7.9**	**18.8-47.9**

Table S6. Demographic characteristics of study sites.

Country *Ethnic group*	**Participants' native language**	**Language in which experiment was conducted**	**Typical participant's English fluency**	**Mass media exposure**	**~% mass media in English**	**Typical participant's education**	**Degree of gender segregation**	**Community or city scale (number of people)**	**Economic mode(s) of participants**
Austria	German	German	moderate	extensive	50	Some college	none	small towns (<5,000) and large cities (>500,000)	Industrial: low and highly skilled
Belgium	Dutch	Dutch	moderate	extensive	75	Some college	none	medium cities (150,000-500,000)	Industrial: low and highly skilled
Brazil	Portuguese	Portuguese	minimal	extensive	<25	8-12 years	none	small (10,000-150,000) and large cities	Small-scale trade industrial: low and highly skilled
China	Chinese	Chinese	moderate	daily	50	Some college	none	large cities	Industrial: low and highly skilled
Denmark	Danish	Danish	moderate	extensive	50	Some college	none	medium cities	Industrial: highly skilled
Egypt	Egyptian Arabic	Egyptian Arabic	moderate	extensive	50	College degree	moderate	large cities	Industrial: highly skilled
India	Kannada	English	moderate	extensive	50	Some college	minimal	medium cities	Industrial: highly skilled
Indonesia	Jakarta dialect of Indonesian	Formal Indonesian	moderate	extensive	75	8-12 years	minimal	large cities	Small-scale trade industrial: low and highly skilled
Italy	Italian	Italian	minimal	extensive	<25	Some college	none	large cities	Industrial: highly skilled
Japan	Japanese	Japanese	minimal	daily	<25	8-12 years	minimal	large cities	Industrial: low and highly skilled
Korea	Korean	Korean	moderate	extensive	25	Some college	none	large cities	Small-scale: trade industrial: low and highly skilled
Namibia *Himba*	Otjiherero	Otjiherero	none	rare	25	1-3 years	minimal	small bands	Small-scale: horticulture, pastoralism

Netherlands	Dutch	Dutch	moderate	extensive	100	Some college	minimal	small and medium cities	Industrial: highly skilled
New Guinea *Sursurunga*	Sursurunga	Neo-Melanesian	minimal	rare	75	4-7 years	minimal	small villages	Small-scale horticulture
New Zealand	English	English	primary language	extensive	100	College degree	none	large cities	Industrial: highly skilled
Peru (rural)	Spanish	Spanish	minimal	extensive	<25	8-12 years	none	small towns	Small-scale horticulture, agriculture, pastoral industrial: low skill
Peru (urban)	Spanish	Spanish	moderate	extensive	50	Some college	none	large cities	Industrial: highly skilled
Singapore	English	English	primary language	extensive	75	Some college	none	large cities	Industrial: low and highly skilled
Slovakia	Slovak	Slovak	minimal	extensive	<25	College degree	none	large towns (5,000-10,000) and small cities	Industrial: highly skilled
South Africa *Zulu*	isiZulu	isiZulu	minimal	occasional	<25	8-12 years	minimal	large villages (200-1,000) and small towns	Small-scale agriculture, pastoralism, trade industrial: low-skill
Spain	Spanish	Spanish	minimal	extensive	<25	Some college	none	small and medium cities	Industrial: low and highly skilled
Tanzania *Hadza*	Hadzane	Hadzane and Swahili	none	rare	<25	1-3 years	moderate	small bands	Hunter-gatherer
Turkey	Turkish	Turkish	moderate	extensive	50	8-12 years	minimal	large cities	Small-scale trade industrial: low and highly skilled
USA	English	English	primary language	extensive	100	Some college	none	large cities	Industrial: highly skilled

Fig. S1. Rating responses (Y-axis) to Question 2 (On a scale of 1 to 7: How much do the speakers like one another?) in each study site broken down by experimental condition (strangers clustered on the left, or friends clustered on the right), and dyad type (male-male, male-female, female-female). Tanzania/Hadza were not asked Question 2 and do not appear here. See Table S3 for means and standard deviation values.

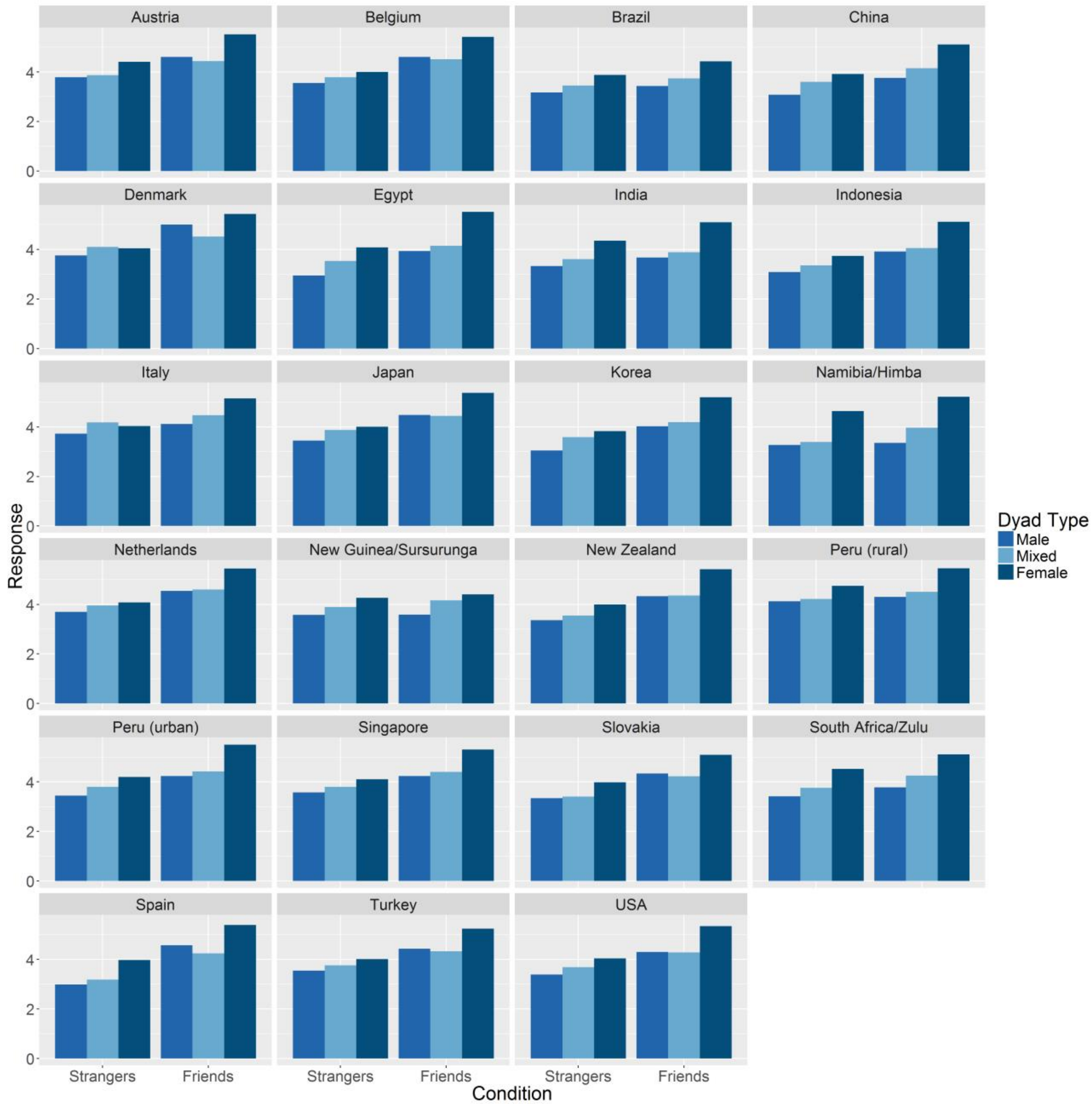
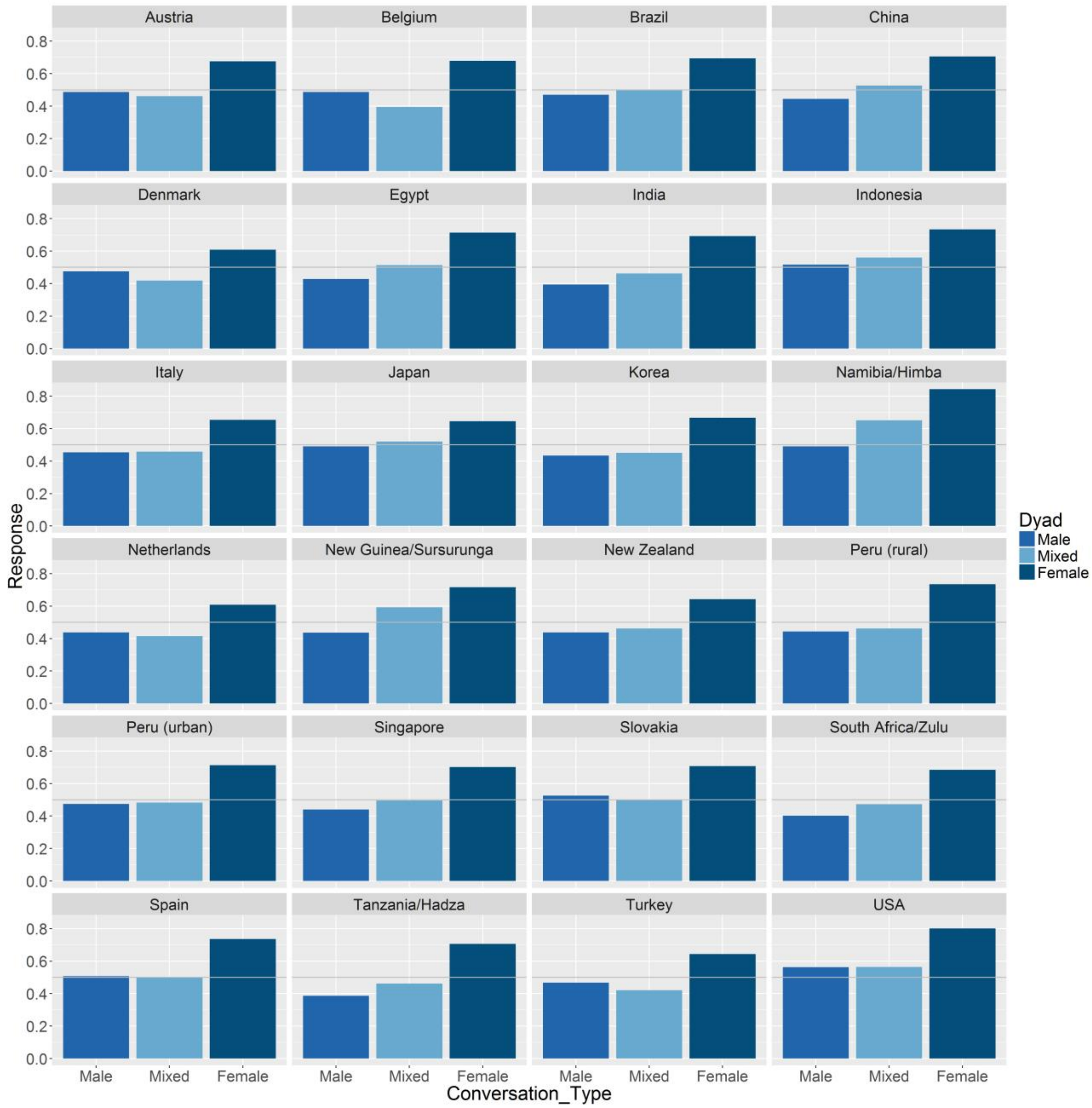


Fig. S2. Mean response rates (0 to 1) on Question 1 across all societies revealing universal bias to respond with answer of "friends" in female-female dyads, but not other dyad types. Y-axis reference line represents unbiased response rate (0.5).



Acoustic analysis of co-laughter segments

Measures

Having demonstrated that participants could accurately judge whether co-laughers are friends or strangers, we then measured a wide range of acoustic features of the laughter to identify which features would best explain the variance in participants' judgments. We examined the 48 co-laughter segments used in the experiment, for a total of 96 individual laughs. Two laughs were excluded (pair 3, participant 1, late laugh; pair 14, participant 2, late laugh) as they did not contain sufficient voicing (i.e., periodic) duration to allow us to automatically assess all of the acoustic features examined. We then analyzed the remaining 94 laughs, 47 of which were produced by friends and 47 of which were produced by strangers.

For each individual laugh within a given audio clip we measured the rate of intervoicing interval (rate of IVI) (1). We first calculated bout duration for each laugh from the onset of visible acoustic energy as viewed in a spectrogram (FFT method, window length: 0.005 s., time steps: 1000, frequency steps: 250, Gaussian window shape, dynamic range: 50 dB) to the offset of energy in the final call, or bout-final inspiratory element. Calls were counted based on audible and visible separated voiced energy. Mean call duration was calculated as total bout duration divided by call number. Mean intervoicing interval (IVI) was calculated as the summed lengths of all unvoiced intervals between calls (i.e., voiced call offset to voice call onset) divided by call number minus one. Unvoiced portions were determined by a lack of formant structure as viewed through a spectrogram with settings described above, and lack of periodicity with standard pitch range values. Finally, rate of IVI was calculated using the following formula:

$$\frac{\left(\frac{(\sum x_i)}{(c-1)}\right)}{\left(\frac{d}{c}\right)}$$

where x_i are the inter-voicing interval values, c is the total call number, and d is the bout duration of the series. This measure captures the averaged rate of unvoiced segments per call across a laugh bout. We also calculated the amount of overlap between co-laughs, that is, the duration for which both laughs can be heard at the same time (co-laugh overlap). Onsets and offsets of each laugh were automatically extracted by individuating the first and last data points in the signal with intensity >30 db, minimizing extraction of possible leakage from each laugh's counterpart. Overlap was calculated as the difference between the earliest offset and the latest onset in each co-laugh pair. Co-laughter between friends and strangers did not differ in overlap as indicated by a mixed model with Overlap as the dependent variable, Familiarity as the fixed factor, and Pair as a random intercept: $\beta = 0.08$, $SE = 0.14$, $t = 0.60$, $P = 0.55$. Nonetheless, in the interests of maximal rigor, we opted to include this variable in the model described below.

Using Praat (2), we extracted fundamental frequency (F_0) (frequency range = 70-400 Hz), and intensity during voiced intervals. F_0 values were converted to a logarithmic scale to approximate perceptual pitch, where after we produced sequences of voiced (presence of pitch) and unvoiced segments. Per each of these measures we calculated traditional descriptive statistics and temporal dynamics measures.

Descriptive statistics. We calculated a) the total and voiced duration of each laugh and the number of voiced and unvoiced segments, and b) standard deviations of pitch and intensity, as well as the mean and standard deviation of length of voiced and unvoiced segments.

Temporal dynamics measures. Traditional descriptive statistics do not capture other crucial aspects of time-series properties such as their regularity over time and the temporal-dependence between successive data points. These properties express the stability and complexity of voice production and have proven particularly useful to assess vocal behavior in a wide variety of contexts (for a review see [3]). To assess these temporal dynamics we employed two non-linear methods: a) Recurrence Quantification Analysis (RQA) of both voiced/unvoiced sequences and pitch (4); and b) Teager–Kaiser energy operator of pitch (5). RQA is a general non-linear time-series analysis tool that quantifies multiple aspects of the temporal stability of a time series, such as how repetitive, noisy, or stationary it is.

Relying on the time series in each laugh (e.g., a sequence of estimated pitch regularly sampled over time), RQA reconstructs the phase space of possible combinations of states and quantifies the structure of recurrence; that is, the number of instances in which the time series displays repeated dynamics, and the characteristics of these repetitions. In order to apply RQA, two steps are necessary: 1) reconstructing the phase space underlying the time series and 2) production of a recurrence plot. The phase space of a time series is an n-dimensional space in which all possible states of a system are represented, so that it is possible to portray the trajectories of the system's behavior, be it periodic (repeatedly crossing the same regions at regular intervals), random, or chaotic. In order to reconstruct the phase space, we applied the time-delay method (6) to each time series. After reconstructing the phase space, we constructed recurrence plots for each time series. Black dots on the plots represent every occasion at which a phase space trajectory goes through approximately the same region in the phase space. In mathematical terms, if we represent the trajectory of a system as

$$\{\vec{x_i}\}_{i=1}^{N}$$

the corresponding recurrence plot is based on the following recurrence matrix:

$$R_{i,j} = \begin{cases} 1: \vec{x_i} \approx \vec{x_j}, \\ 0: \vec{x_i} \not\approx \vec{x_j}, \end{cases} i, j = 1, \ldots, N$$

where N is the number of considered states of the system and $\vec{x_i} \approx \vec{x_j}$ indicates that the two states are equal up to an error (or distance) ε. Note that this ε is essential in the case of continuous variables (as in F_0) as systems often do not recur exactly, but only approximately revisit states. To statistically analyze differences in laughs, we performed Recurrence Quantification Analysis (RQA) on the recurrence plots. RQA provides several indices quantifying the structure and complexity of dynamical systems from recurrence plots (4). This makes it possible to statistically compare different dynamic systems (e.g., different dyads) in terms of their dynamics such as the stability, structure, and complexity in the behavior of the system. In particular we analyzed:

Amount of repetition: The percentage of values that recur (are repeated) in the time series independently of the lag (recurrence rate, RR).

$$RR(\varepsilon) = \frac{1}{N^2} \Sigma_{i,j=1}^{N} R_{i,j}(\varepsilon)$$

Stability of repetition: articulated in:

Average length of sequences repeated (L)

$$L = \frac{\sum_{l=lmin}^{N} l\, P(l)}{\sum_{l=lmin}^{N} P(l)}$$

Length of longest repeated sequence (LMAX)

$$LMAX = \max(\{l_i\}_{i=1}^{N_l})$$

For more details about these indices see (4).

The Teager–Kaiser energy operator (TKEO) has been widely employed to quantify jitter and shimmer; that is, perturbations in the regular cycles of pitch and intensity, respectively, which often characterize situations of stress and arousal, and are impacted by the ability to control the speech production system (5). TKEO is calculated as

$$\psi(x_n) = x_n^2 - x_{n+1} \cdot x_{n-1}$$

where the subscript n denotes the n^{th} entry of the vector x (in our case, the time series of pitch). We computed the mean, standard deviation and 5th, 25th, 75th and 95^{th} percentile values of TKEO.

Overall, this resulted in 34 features for each individual laugh (Table S7).

Besides individual features in each laugh, we were also interested in the relationship between laughs in a co-laughter dyad: do co-laughs between friends share features that co-laughs between strangers do not? Therefore, we prepared a second dataset by calculating the difference in values across all features for each pair of laughs. This yielded a dataset of 34 features for each of the 46 pairs of laughs for which we had acoustic features of both of the individual laughs making up the pair. These two datasets were used separately to assess which acoustic features listeners might employ when judging whether a co-laugh was between friends or strangers. We call this measure the *Friends Ratio* (FR), defined as the overall likelihood of each single laugh being part of a co-laugh segment produced between individuals identified by participants as being friends. In order to examine cross-cultural reliability, we then employed the overall model to predict within-cultures FR and assessed the amount of variance explained through Adjusted R^2. All acoustic features were linearly transformed on a scale from 0 to 1 for better performance in the feature selection process.

Analysis and machine learning process

Feature selection. The previously described process produces a large set of features, exemplifying what is commonly termed the curse of dimensionality. In other words, the presence of a large number of features makes the statistical models both difficult to interpret and at risk of overfitting, producing results that are not generalizable. To address this, we used a common algorithm to select a parsimonious subset of features: the Elastic Net extension of the LASSO (7), which could in principle reduce overall accuracy, but increases the interpretability and generalizability of the results, that is, the ability to accurately describe new laughs with characteristics similar to the laughs in the current study.

Statistical models. In order to assess the overall model relying on the selected features, we used a 5-fold cross-validated multiple regression model to reconstruct the participants' likelihood of judging a given dyad of co-laughers to be friends (FR). The dataset was divided into 5 subsets (or folds) each containing a non-overlapping fifth of the pairs of co-laughers. A combination of 4 folds was used for feature selection and model fitting. The model was then assessed on the remaining fold. This procedure was repeated for all four possible combinations of folds. In this way the accuracy of the model was assessed only on data on which it had not been trained. We repeated the cross-validation process a total of 100 times, randomly permuting the data before splitting into training and testing subsets to ensure stability of the results across different random splits in 5 folds.

Table S7. 34 extracted features in acoustic analysis.

Voiced / Unvoiced segments	Pitch	Intensity
Co-laugh overlap	Recurrence Rate (RR)	Recurrence Rate (RR)
Total duration of the laugh	Average length of recurrent sequence (L)	Average length of recurrent sequence (L)
Voiced duration	Maximum length of recurrent sequence (LMAX)	Maximum length of recurrent sequence (LMAX)
Unvoiced duration	Pitch SD	Intensity SD
Number of unvoiced segments	Average TKEO	Average TKEO
Average length of unvoiced segments	SD of TKEO	SD of TKEO
Average length of voiced segments	5^{th} percentile of TKEO	5^{th} percentile of TKEO
SD of length of unvoiced segments	25^{th} percentile of TKEO	25^{th} percentile of TKEO
SD of length of voiced segments	75^{th} percentile of TKEO	75^{th} percentile of TKEO
Mean intervoicing interval	95^{th} percentile of TKEO	95^{th} percentile of TKEO
Mean call duration		
Recurrence Rate (RR)		
Average length of recurrent sequence (L)		
Maximum length of recurrent sequence (LMAX)		

Acoustic analysis results

The statistical model employing acoustic features of individual laughs was able to statistically predict FR to a high degree: R^2 = 0.43 (CI: 0.29 0.57), *Adjusted* R^2 = 0.42 (CI: 0.28 0.56), p = 0.0001. Features selected were: 1) Mean call duration; 2) Pitch average TKEO (mean jitter); 3) Pitch SD TKEO (SD jitter); 4) Intensity 5th percentile TKEO (5th PCTL shimmer). See Table 2 in main text.

In summary, the findings suggest that individual laughs that had 1) shorter average call duration, 2) less regular pitch cycles, 3) less variation in pitch cycle regularity, and 4) less regular intensity cycles were more likely to be rated as having been produced by friends. The model remains quite consistent across cultures as it explains a significant portion of the variance in the FR within each culture. See Table S8 for R^2 and Adjusted R^2 values in each culture. Figure 3 in main text displays a scatterplot showing the correlation between participants' friend response across all cultures and predicted values using acoustic features selected by the statistical model.

Table S8. Acoustic feature selection model performance predicting friend response across 24 cultures.

Group	**Region**	R^2	**Adjusted** R^2
South Africa/Zulu	Africa	0.37	0.37
Egypt	Africa	0.34	0.33
Namibia/Himba	Africa	0.05	0.04
Tanzania/Hadza	Africa	0.30	0.29
Singapore	Asia	0.38	0.37
Korea	Asia	0.36	0.36
Japan	Asia	0.16	0.15
India	Asia	0.32	0.31
China	Asia	0.32	0.31
Indonesia	Asia	0.29	0.28
Denmark	Europe	0.20	0.19
Netherlands	Europe	0.19	0.18
Belgium	Europe	0.36	0.35
Slovakia	Europe	0.38	0.38
Turkey	Europe	0.21	0.20
Spain	Europe	0.41	0.41
Austria	Europe	0.38	0.37
Italy	Europe	0.35	0.34
USA	N. America	0.32	0.32
New Guinea/Sursurunga	Oceania	0.21	0.20
New Zealand	Oceania	0.34	0.33
Peru (urban)	S. America	0.38	0.37
Peru (rural)	S. America	0.50	0.49
Brazil	S. America	0.33	0.32

Responses provided by Himba (Namibian) participants constitute an exception to the model's ability to successfully explain a large proportion of the variance in FR across cultures, as the model explained only 5% of the variance in FR responses in this subsample. A separate model was trained on the Himba sample exclusively that explained slightly increased variance in their FR responses: Adjusted $R^2 = 0.08$ (CI: 0.04 0.14), $p = 0.04$. Features selected were: 1) lower mean call duration, and 2) less variability in pitch cycle regularity. See Table S9.

The Himba participants exhibited a stronger bias to judge co-laughers to be friends (69%) than is true in any other culture. B. Scelza, the co-author who collected these data, observed that some participants were occasionally confused by the question of "friends versus strangers," as this dichotomy does not allow for the identification of pairs of laughers who know each other but are not friends. However, closer analysis of the relationship between Himba participants' responses to Question 1 (friends versus strangers) and their responses to Question 2 (how well the pair liked one another) did not reveal an unusual pattern, suggesting that Himba participants interpreted Question 1 in a similar manner to participants in the other cultures studied.

Table S9. Sample coefficients from one run of the 5-fold cross-validated model on Himba sample.

Predictor	Beta (SE) Fold1	Beta (SE) Fold2	Beta (SE) Fold3	Beta (SE) Fold4	Beta (SE) Fold5
Intercept	0.890 (0.051)	0.944 (0.046)	0.873 (0.036)	0.885 (0.038)	0.924 (0.041)
Mean call duration	-0.336 (0.108)	-0.382 (0.103)	-0.381 (0.088)	-0.218 (0.093)	-0.570 (0.106)
Pitch TKEO SD	-0.391 (0.110)	-0.601 (0.119)	-0.298 (0.098)	-0.298 (0.103)	-0.242 (0.100)

The model employing the difference in acoustic features between co-laughers' laughs was not able to statistically predict FR to any significant degree. One limitation in this analysis for assessing coordination between co-laughers is the circumscribed overall length of the co-laughs (approximately 1 s.). An evolved signal of close affiliation would likely reveal the capacity for coordinated production, and this may be difficult to assess in short bouts of co-laughter isolated from the normal discursive context wherein bouts occur repeatedly. Future work should therefore investigate longer bouts of co-laughter to examine whether affiliated speakers produce coordinated laughter in a way that strangers do not, thus constituting more solid evidence that a signaling system is in place. At present, the results confirm, however, an available cue of affiliation that can be employed with quite limited information to make accurate judgments about the relationship between pairs of friends in dialogue.

Cultural and linguistic demographic dimensions

In principle, the extent to which third-party listeners can determine the relationship between co-laughers could be a function of many aspects of the listeners. For example, if the language spoken substantially influences the form of laughter, then, given that the stimuli were all generated by Californian speakers of American English, we might expect that listeners' familiarity with English would influence the accuracy of their judgments in this regard. Familiarity plausibly derives not only from the language spoken by the listener, but also from exposure to mass media in which English is used. More broadly, cultural similarity independent of familiarity with English could play a role. Likewise, a wide variety of studies in psychology reveal that individuals who are highly educated resemble one another across cultures (8), hence listeners' level of education might plausibly contribute to their ability to accurately discern the relationships obtaining between the undergraduate students whose laughter constituted the stimuli. Additionally, because participants were asked to make assessments of stimuli provided by dyads of each sex, and by mixed-sex dyads, the degree of gender segregation characteristic of a society could conceivably influence participants' accuracy. If experience plays a role in the ability to judge the nature of relationships on the basis of laughter, then individuals from highly gender-segregated societies would plausibly have less experience with mixed-sex dyads and dyads of the opposite sex, and thus might be expected to be less accurate. Similarly, if experience is critical, then participants from societies organized on the basis of small groups would necessarily have less experience with the range of idiosyncratic laughter styles possible than would participants from large cities, and thus could be expected to be less accurate. The manner in which an experiment is implemented could also have an effect on participants' responses, including whether instructions were read aloud to the participant by the researcher, or read visually by the participant on a computer screen. With these considerations in mind, the investigators responsible for each study site in this project estimated the features of their participants relevant to the above considerations; said data were not collected directly from the participants themselves, in part because, for some of the samples employed, questions concerning these matters could have damaged investigator rapport with the participants and/or disrupted the research process. Table S6 presents a summary of this information. Importantly, as noted in the text, despite the substantial multidimensional variation across study sites summarized in Table S6, neither the identity of the participants nor the location of the study sites accounted for substantial variance in the key dependent measures, thus underscoring the universality of the cues of affiliation presented by co-laughter.

Complete text of English instructions for participants

The following text was used in the computerized experiment for English speakers, or as the basis for translation. The numbering below denotes separate screen presentation in the SuperLab experiment.

1) Welcome to the laughter and friends study. In this experiment we will have you listen to recordings of people in pairs laughing together in a conversation, and then answer two questions about each recording.
2) Some of the pairs of people were friends at the time of the recording, and others were complete strangers who were meeting for the first time. We will ask whether you think the people laughing together were friends or were strangers, and then we will ask you how well you think the people liked each other.
3) The recordings of people laughing are very short, and do not include any conversation. Before we begin with the actual study, you will be able to practice with one recording so that you will be familiar with the procedure.
4) Press the space bar when you are ready to begin the practice session.
5) When you are ready, press the space bar to hear the practice recording.
6) Do you think these people laughing were friends at the time of the conversation, or were they strangers who had just met for the first time? Press 1 if you think that the people in the recording were friends and 0 if you think that the people in the recording were strangers.
7) How much do you think these people liked each other? Provide your estimate using this s scale, where 1 means "Not at all," 4 means "Somewhat," and 7 means "Very much."
8) When you are ready, press the space bar to hear the recording. If you have any questions, please ask the experimenter. If you do not have any questions, press the enter key and the experiment will begin.
9) You have now listened to all of the recordings. Thank you for your participation. Please tell the experimenter that you are finished.

Auxiliary Experiments

To further explore the perception of our co-laughter stimuli, we ran two additional studies using the same participant pool as the US sample reported in the main text. Our aim in these studies was to better assess the extent to which judgments of the individual laughs making up the co-laugh pairs were responsible for our cross-cultural findings.

Study 1 – Perceiving arousal and valence in individual laughs. Judgments of whether people laughing together are friends or strangers could be mediated by perceivable affective characteristics in the individual laughs making up a given co-laugh pair. We investigated how percepts of emotion in the laughs making up our stimulus set were related to judgment patterns of the co-laugh pairs reported in the main text.

Research examining the perception of vocal emotion, including other studies of laughter (e.g., 9), often use a dimensional approach to emotion (10) that conceptualizes affect as working on at least two independent dimensions: arousal and valence. The dimension of arousal works on a scale of low activity (e.g., calm) to high activity (e.g., excited), and valence works on a scale of positive (e.g., happy) to negative (e.g., sad). All emotions, in theory, can be minimally described in this two-dimensional space. For example, "fear" can be described as high arousal with negative valence and "sadness" can be described as low arousal with negative valence.

If individual laughs are perceived as being high in arousal and positive in valence, participants could then decide that the laughter is likely being produced in the presence of a friend. Conversely, low arousal and negative valence (or less positive valence) could be used as a cue that the laughter occurred with a stranger. We expected, therefore, that ratings of arousal and valence would be positively related to the likelihood that our judges across the 24 study sites identified a given pair containing the individual laughs as occurring between friends.

Methods and Results. We isolated the individual laughs from all co-laugh pairs used in our main study, and then presented these individual 96 laughs (2 laughs each from 48 pairs) to 24 participants (15 female, and 9 male) who received credit in an introductory communication course at UCLA. Judges listened to all 96 laughs presented in random order, and then rated the laughs for arousal and valence on a scale of -5 (not arousing, very negative) to +5 (very arousing, very positive). Both terms were defined for the participants. Figure S3 shows the mean ratings of arousal and valence across the three dyad types from which the laughs were taken (male, mixed, and female pairs).

Variation in arousal and valence judgments was explained by speaker familiarity (Arousal: $F(1, 44) = 27.93$, $P < 0.0001$, $\eta^2 = 0.56$; Valence: $F(1, 44) = 22.10$, $P < 0.0001$, $\eta^2 = 0.50$) as well as dyad type (Arousal: $F(2, 44) = 13.93$, $P < 0.0001$, $\eta^2 = 0.39$; Valence: $F(2, 44) = 13.62$, $P < 0.0001$, $\eta^2 = 0.38$). Individual laughs taken from co-laugh pairs between friends were judged as having higher arousal than individual laughs produced in stranger dyads, and laughs between female friends were rated higher on both dimensions. Familiarity and dyad type interacted on both affective dimensions, however, with ratings of arousal and valence both higher on female friends than other dyad/familiarity combinations (Arousal: $F(2, 44) = 5.84$, $P < 0.01$, $\eta^2 = 0.21$; Valence: $F(2, 44) = 4.74$, $P < 0.05$, $\eta^2 = 0.18$).

As described earlier, individual laughs were taken from co-laugh pairs that were judged by our 966 participants worldwide on whether they were from pairs of friends or strangers. For each individual laugh that made up a pair, ratings of arousal and valence were averaged across our 24

US judges for all 48 laughter pairs. The likelihood of a laugh being judged (when in a co-laugh pair) as being between friends was strongly positively associated with averaged ratings of arousal ($r = 0.87$, $P < 0.0001$) and valence ($r = 0.82$, $P < 0.0001$). See Figure S4 for both correlations.

Fig. S3. Average ratings of arousal and valence for individual laughs across three dyad types from which the laughs originated (as co-laugh pairs).

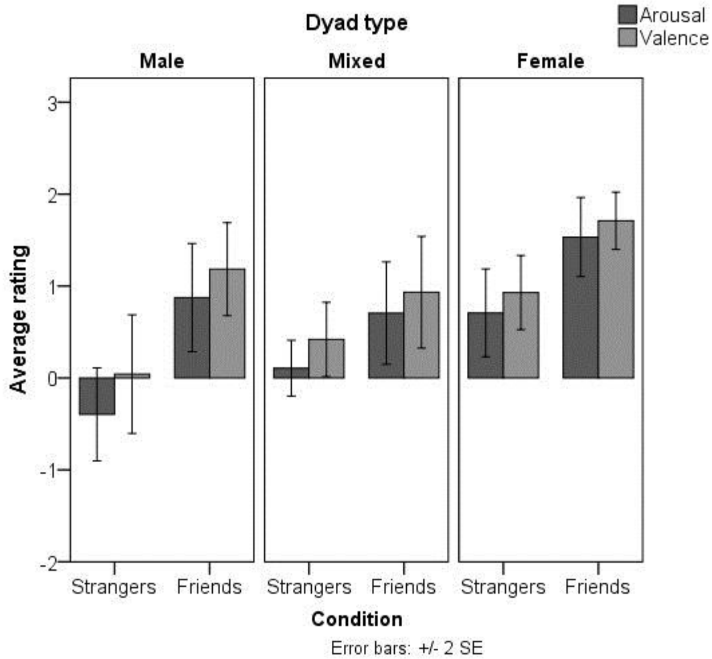
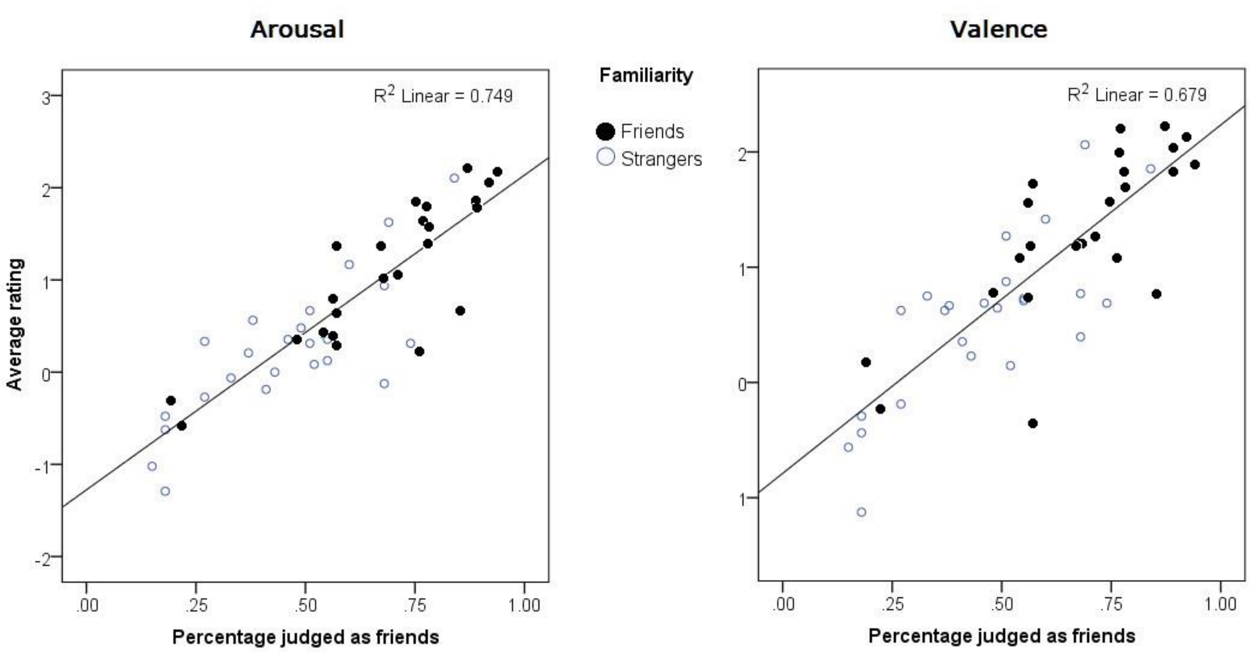


Fig. S4. Scatterplots showing averaged ratings of arousal and valence of individual laughs associated with percentage judged as friends when heard as co-laugh pairs.



Discussion. When US participants listened to individual laughs taken from co-laugh pairs between friends and strangers and asked to rate them on two affective dimensions, judgments patterned in a manner that closely mirrored our cross-cultural evaluations of whether pairs of laughers were friends or strangers. This finding suggests that listeners are tapping into affective characteristics of individuals' laughs when deciding whether two people are established friends or newly acquainted strangers. Greater arousal and more positive valence was detected in laughter between friends, reflecting the emotional signaling that often occurs between familiar people in a conversation. Moreover, we found that, just as in our large cross-cultural sample, female friends exhibited the most detectable signals. In the current study, judges could detect increased arousal and more positive valence in laughter between female friends, and across all cultures in the main study, co-laughter pairs between female friends were more recognizable than other dyad/familiarity combinations (and more likely to elicit that judgment). Finally, the arousal and valence ratings were very highly positively correlated with judgments in our main study of whether the co-laughers were friends, strongly suggesting that the affective dimensions are informing listeners about affiliation status.

Study 2 – Judging affiliation in shuffled co-laughter pairs. If people's judgments of affiliative status between co-laughing individuals are driven by individual laugh characteristics, then judges should remain accurate when individual laughs are presented in artificially created pairs of "friends." Individual laugh characteristics might be driving people judgments to the extent that actual friend pairs are no more likely to be judged as such as constructed pairs consisting of individual laughs produced with other people who are friends. To explore this possibility, we created artificial friend co-laughter samples from our stimulus materials, and repeated the experiment.

Our acoustic analysis examining within-pair co-laughter differences failed to reveal any link between the laughter production dynamics (see acoustic analysis) but it is still possible that some production connection between the co-laughers was perceivable. We expected that artificial friend co-laughter would also be judged as originating from friends, but we also anticipated that differences in the judgment patterns from real friend pairs would reveal important aspects of the laughter that help judges make accurate evaluations. If temporally dynamic properties are detectable in such a short time frame, we might expect artificially familiar co-laugh pairs to not be readily identified as between friends. But if individual laughter characteristics are primarily responsible for listeners' judgments, then we should see a similar pattern of evaluations in artificial pairings, since the pairs would consist of people laughing with friends (albeit not the laughers' partner in the co-laugh the participants hear).

Methods and Results. Using the 48 co-laughter stimuli, we switched individual laughs within familiarity (24 friend pairs, 24 stranger pairs) and dyad type (6 male pairs, 8 mixed pairs, 10 female pairs) categories. For any given laugh, it was paired with another laugh in the same category ($2 \times 3 = 6$ categories). The new artificial pairs, then, were still representative of their original category, except the specific individuals had not produced their laughs together, and none of the pairs represented actual friends that had engaged in a real conversation.

The 48 new laugh co-pairs were presented to 37 participants (25 female, and 12 male) who received credit in an introductory communication course at UCLA. The procedure was identical to the experiment described in the main text. Participants listened to the 48 new co-pairs in random order and answered two questions for each trial. Overall accuracy (measured as identifying a constructed co-pair made up of individual laughs produced between two friends as “friends” and two strangers as “strangers”) was 65% (SD = 0.48) which was better than chance (z = 11.38, $p < 0.0001$). Figure S5 shows the rates of correct judgments across the six condition and dyad type combinations (Question 1) and Figure S6 shows the averaged rating data (Question 2). We again used a model comparison approach (GLMMs with Laplace approximation with effects on model fit measured by AIC) combining variables as done in the main experiment with the obvious exception of the random factor of culture. The best fitting model for both questions was similar to the original model across our 24 societies, with the exception that participant sex did not lower AIC as it did in the main study with the original co-pairs (i.e., female participants were more accurate in original study). But condition (friends versus strangers) as well as dyad type (male, mixed, and female) both revealed main effects, and these variables interacted similarly as in the main study. See Table S9 for best fitting model values for question 1, and Table S10 for question 2.

Discussion. When the individual laughs making up our original co-pairs were recombined into new co-pairs within the same categories, a new group of judges was able to accurately recognize whether the co-pairs were composed of laughs between friends or strangers. While the judges were asked about the pair they heard, and not about the individuals themselves, listeners clearly had no trouble assuming the novel pairs were friends. The pattern of accuracy, while not exactly the same, was very similar to the cross-cultural experiment reported in the main text, as evidenced by a similar best-fitting model for the forced-choice data. The same pattern was found for the rating question as well, with female friends thought to like each other more than the other categories. These data are consistent with the notion that individual laugh characteristics were important for our cross cultural judges in determining whether co-laughing people were friends or strangers.

Fig. S5. Judgments of affiliation in artificially paired co-laughs across familiarity conditions and dyad types.

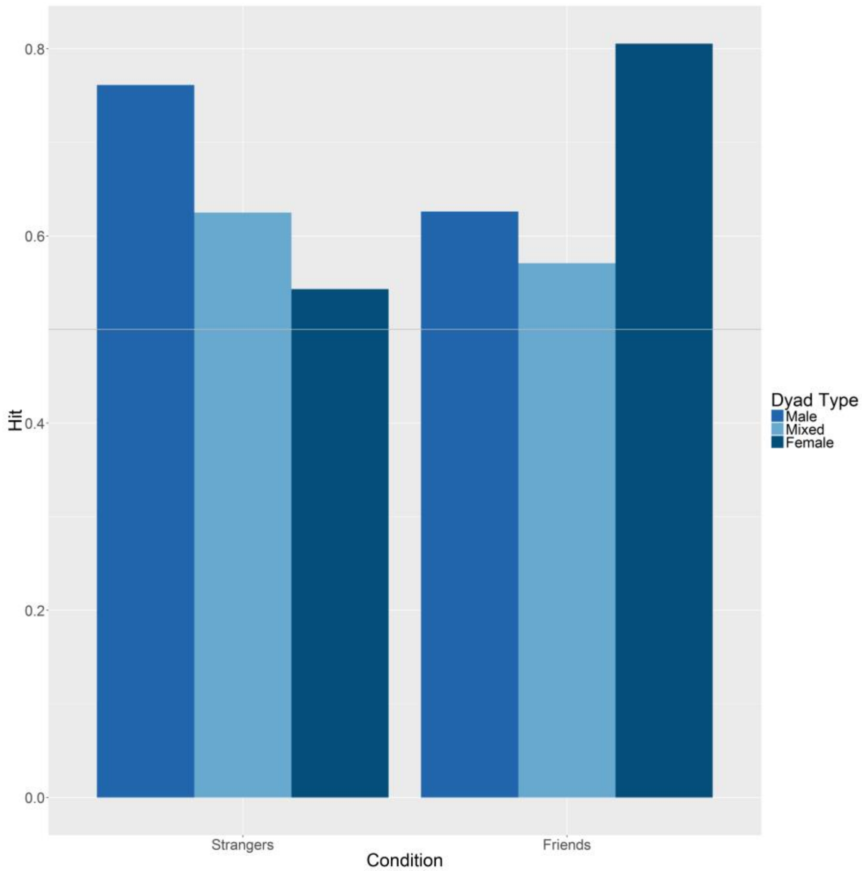


Table S9. Best-fit model for Question 1 (Accuracy of judgments of relationship obtaining between co-laughers).

Random effect			**Fixed effects**				
Factor	***Variance***	***STD***	***Factor***	***Estimate***	***SE***	*z*	*Pr(>|z|)*
Subject	0.03057	0.1749					
			Intercept	1.1676	0.1606	7.269	3.61e-13 **
			Condition	-0.6481	0.2105	-3.079	0.00208 *
			ConvType1	-0.6530	0.1986	-3.288	0.00101 *
			ConvType2	-0.9929	0.1896	-5.238	1.62e-07 **
			Condition × ConvType1	0.4214	0.2696	1.563	0.11815
			Condition × ConvType2	1.9032	0.2697	7.057	1.70e-12 **

Fig. S6. Ratings of liking in artificially paired co-laughs across familiarity conditions and dyad types.

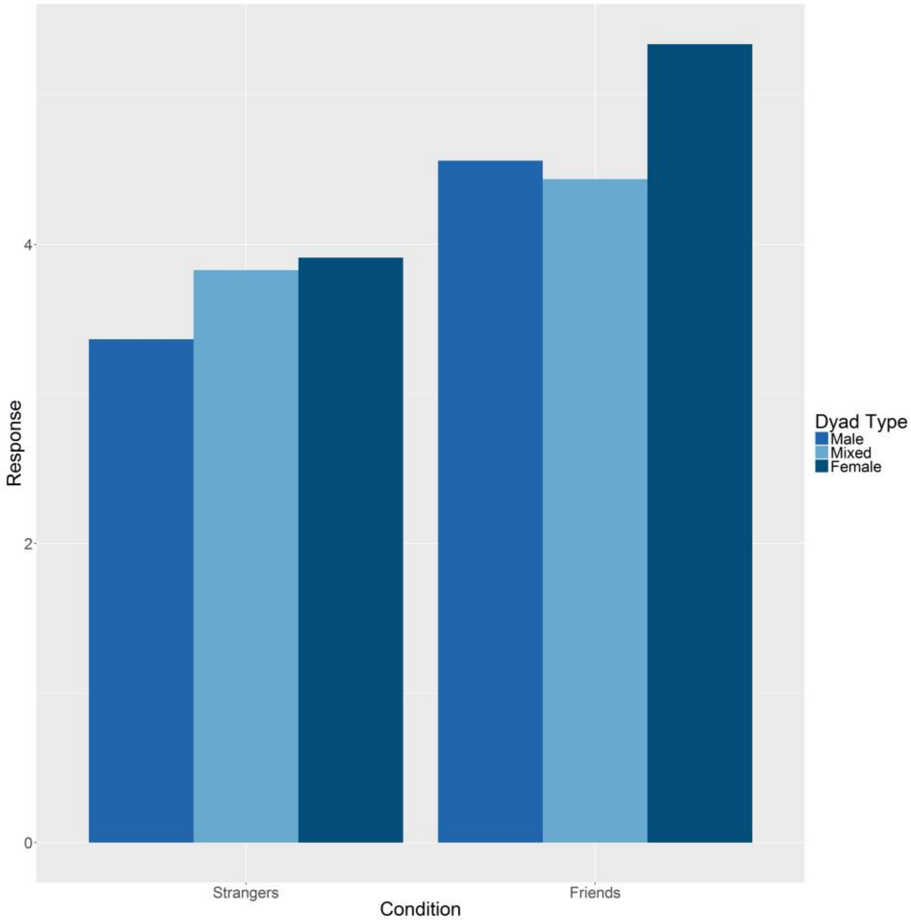


Table S10. Best-fit model for Question 2 (Ratings of liking between co-laughers).

Fixed factors	Random factor	Estimate	SE	t	Variance	SD
(Intercept)		3.3649	0.1143	29.438		
Condition1		1.1937	0.1334	8.948		
ConvType1		0.4628	0.1248	3.709		
ConvType2		0.5459	0.1193	4.576		
Condition x ConvType1		-0.5856	0.1765	-3.318		
Condition x ConvType2		0.2333	0.1687	1.383		
	Participant				0.1542	0.3927

References

1. Bryant GA, Aktipis CA (2014) The animal nature of spontaneous human laughter. *Evol Hum Behav* 35(4): 327-335.
2. Boersma P, Weenink D (2014) Praat: Doing phonetics by computer [Computer program] Version 54, retrieved 4 October 2014 from http://wwwpraatorg/
3. Fusaroli R, Konvalinka L, Wallot S (2014) Analyzing social interactions: Promises and challenges of cross recurrence quantification analysis. *Springer Proceedings in Mathematics & Statistics*.
4. Marwan N, Carmen Romano M, Thiel M, Kurths (2007) Recurrence plots for the analysis of complex systems *Phys Rep* 438(5): 237-329.
5. Tsanas A, Little MA, McSharry PE, Ramig LO (2011) Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson's disease symptom severity. *J Roy Soc Interface*, 8(59): 842-855.
6. Takens F (1981) Detecting strange attractors in turbulence. *Dynamical systems and turbulence, Warwick 1980* (Springer, Berlin), pp 366-381.
7. Zou H, Hastie T (2005) Regularization and variable selection via the elastic net. *J Roy Stat Soc B* 67, 301-320.
8. Henrich J, Heine SJ, Norenzayan A (2010) The weirdest people in the world? *Behav Brain Sci* 33(2-3): 61–83.
9. Sauter DA, Eisner F, Ekman P, Scott SK (2010) Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proc Natl Acad Sci* 107(6): 2408–2412.
10. Russell JA (1980) A circumplex model of affect. *J Pers Soc Psychol* 39(6): 1161-1178.

Audio and video samples

Six audio (.wav format) examples of laughter types, **s**ix sample stimulus files, and one video file showing Hadza participants are provided online (13 total). All laughter recorded by GAB except S2.

Audio S1:	Normal laugh with periodic structure and egressive airflow (human_laugh.wav)
Audio S2:	Chimpanzee "laugh" (*Pan troglodytes*) with noisy, aperiodic structure and alternating airflow (chimp_laugh.wav) (Courtesy of Robert Provine).
Audio S3:	Voiced laugh (voiced_laugh.wav)
Audio S4:	Unvoiced laugh (unvoiced_laugh.wav)
Audio S5:	Spontaneous laughter (spontaneous_laugh.wav)
Audio S6:	Volitional laughter (volitional_laugh.wav)
Audio S7:	Friends, male-male pair (friends_mm.wav)
Audio S8:	Friends, male-female pair (friends_mf.wav)
Audio S9:	Friends, female-female pair (friends_ff.wav)
Audio S10:	Strangers, male-male pair (strangers_mm.wav)
Audio S11:	Strangers, male-female pair (strangers_mf.wav)
Audio S12:	Strangers, female-female pair (strangers_ff.wav)
Video S1:	Video of Hadza participant hearing laughter samples (Courtesy of Dr. Coren Apicella).