UNIVERSITY OF HELSINKI

https://helda.helsinki.fi

# Peer evaluation can reliably measure local knowledge

**Reyes-García, Victoria; Díaz-Reviriego, Isabel; Duda, Romain; Fernández-Llamazares, Álvaro; Gallois, Sandrine ...**

**2016**
SAGE Publications Inc.
http://hdl.handle.net/10138/173801
Reyes-García, V, Díaz-Reviriego, I, Duda, R, Fernández-Llamazares, Á, Gallois, S, Guèze, M, Napitupulu, L & Pyhälä, A 2016, 'Peer evaluation can reliably measure local knowledge', Field Methods, vol. 28, no. 4, pp. 1-18. https://doi.org/10.1177/1525822X16629912

Downloaded from Helda, University of Helsinki institutional repository. https://helda.helsinki.fi
This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.
Please cite the original version.

*Article*

Field Methods
1-18
© The Author(s) 2016
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1525822X16629912
fmx.sagepub.com



# Peer Evaluation Can Reliably Measure Local Knowledge

**Victoria Reyes-García[1,2], Isabel Díaz-Reviriego[2], Romain Duda[2], Álvaro Fernández-Llamazares[2,3], Sandrine Gallois[2], Maximilien Guèze[2], Lucentezza Napitupulu[2], and Aili Pyhälä[2,3]**

## Abstract

We assess the consistency of measures of individual local ecological knowledge obtained through peer evaluation against three standard measures: identification tasks, structured questionnaires, and self-reported skills questionnaires. We collected ethnographic information among the Baka (Congo), the Punan (Borneo), and the Tsimane' (Amazon) to design site-specific but comparable tasks to measure medicinal plant and hunting knowledge. Scores derived from peer ratings correlate with scores of identification tasks and self-reported skills questionnaires. The higher the number of people rating a subject, the larger the association. Associations were larger for the full sample than for subsamples with high and low rating scores. Peer evaluation can provide a more affordable method in terms

[1] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain
[2] Institut de Ciència i Tecnologia Ambientals, Universitat Autònoma de Barcelona, Barcelona, Spain
[3] Metapopulation Research Centre (MRC), Department of Biosciences, University of Helsinki, Helsinki, Finland

**Corresponding Author:**
Victoria Reyes-García, Institució Catalana de Recerca i Estudis Avançats (ICREA), 08001 Barcelona, Spain.
Email: victoria.reyes@uab.cat

Downloaded from fmx.sagepub.com by guest on February 27, 2016

of difficulty, time, and budget to study intracultural variation of knowledge, provided that researchers (1) do not aim to describe local knowledge; (2) select culturally recognized domains of knowledge; and (3) use a large and diverse (age, sex, and kinship) group of evaluators.

## Introduction

The interest in the potential benefits of traditional ecological knowledge, *sensu* (see Berkes et al. [2000] for a definition), has sparked a growing amount of research in traditional, local, and indigenous knowledge systems. Much of this research has provided detailed descriptions of how societies interact with elements in their surrounding environment (Berkes 1999; Posey 1999). Some other research has taken an hypotheses-driven approach to examine the patterns that model the intracultural distribution of knowledge (Boster 1986; Reyes-García et al. 2005), the modes of knowledge transmission (Demps et al. 2012; Hewlett et al. 2011), the drivers and rate of loss of knowledge (Gómez-Baggethun et al. 2010), or the individual benefits of local knowledge (McDade et al. 2007).

Given the rising importance of the hypotheses-driven approach, the accurate measurement of individual levels of local knowledge is of increasing concern (Kightley et al. 2013; Zent and Maffi 2010). Researchers differentiate between theoretical knowledge and skills: One question is to know the potential medicinal use of a plant; another is to be able to find the plant in the forest, collect it, and prepare it for use (Reyes-García, Martí-Sanz, et al. 2007). Several methods have been proposed to collect and analyze measures of individual levels of knowledge and skills (see Reyes-García, Martí-Sanz, et al. [2007] and Zent and Maffi [2010] for reviews). For example, to measure knowledge, authors have used different cognitive tasks, including free lists (Atran et al. 2002), paired comparisons (Reyes-García et al. 2004), or multiple choice (Reyes-García, Vadez, et al. 2007); to measure skills, authors have used transect surveys (Zarger and Stepp 2004), species identification (Begossi 1996), skill tests (Demps et al. 2012), and self-reports (Ticktin and Johns 2002).

A major concern has been to identify the "correct" answer with which to evaluate local knowledge, as researchers inquiring about local knowledge do not necessarily have enough information to evaluate the accuracy of answers provided. To overcome the problem, researchers have compared the answers from a given informant with the answers provided by other informants in the same society, a theoretical and methodological approach

Downloaded from fmx.sagepub.com by guest on February 27, 2016

known as "cultural consensus," the intuition being that informants' agreement stands for cultural knowledge (Romney et al. 1986). Other researchers have evaluated informants' answers by comparing them against data collected by scientists (Huntington 2000). Researchers have also tried to measure individual levels of local knowledge by using indices that mimic those used in biology, such as the species richness index (Anadón et al. 2009; Begossi 1996). A common limitation to the three approaches is that they largely depend on the criteria of the researcher, often a relatively naive outsider lacking the necessary cultural information to capture all the nuances of a given local knowledge system.

An alternative approach to evaluate an individual's level of knowledge would be to ask about the individual's knowledge from other people in the same society. Peers might be better evaluators than researchers as they share with the subject the overall corpus of knowledge being evaluated and hence are in a better position to provide a more comprehensive and holistic view of the subject's level of knowledge. Additionally, while researchers can only capture knowledge at a single point in time, peers can base their judgment on long-term evaluations.

In itself, the idea to ask peers to evaluate a person's local knowledge is not new. Key informants are often selected by asking peers about "knowledgeable people" (Davis and Wagner 2003). Previous works have also systematized peer evaluations to assess individual levels of local knowledge. For example, Davis and Wagner (2003) used a structured survey technique based on peer recommendations to find out the most knowledgeable fishers; Kightley and colleagues (2013) inferred subjects' skills by asking peers to rank the quality of a set of items made by other villagers; and Demps and colleagues (2012) used a panel of peers to assess honey collection skills based on informants' performance on a series of games (i.e., tree climbing, torch making). While presenting important innovations, previous work has not assessed the accuracy of peer evaluations, which can—theoretically—be challenged in several ways.

First, evaluations might be biased by the evaluator–subject relation. Kin or kith might receive better evaluations than unrelated subjects. Conversely, evaluators might be less familiar with the knowledge of subjects outside their network, which might translate in ratings inaccuracy. Furthermore, in some contexts, it might be culturally inappropriate to publicly evaluate friends or family members, or evaluations might be affected by other criteria, such as political power or gender stereotypes.

Second, expertise might be easier to evaluate in domains that are more culturally important, such as domains related to criteria that determine mate

Downloaded from fmx.sagepub.com by guest on February 27, 2016

selection (Pillsworth 2008), or in domains that result in outputs that are easier to evaluate than others. Alternatively, evaluators might find it difficult to assess individual levels of knowledge for activities that are performed in a collective way or for activities that are culturally confined to the private sphere.

Third, peer evaluation might yield different results at different levels of knowledge. Thus, while it might be relatively easy to assess high levels of knowledge (i.e., to identify experts), peers might be less accurate in differentiating between people with average or low levels of knowledge.

Given these potential limitations, the goal of this work is to assess the consistency of measures of individual local ecological knowledge obtained through peer evaluation against results obtained with methods previously used to measure individual levels of local knowledge.

## Methodological Approach

We collected data in three indigenous, small-scale, subsistence-based societies with little involvement in market economies, school-based education, or modern health care systems: the Baka (Congo Basin), the Punan (Borneo), and the Tsimane' (Amazon). Six researchers conducted 18 months of fieldwork each (two in each indigenous group; each one in a different village). Researchers devoted the first five months to collecting contextual and ethnographic information and the following 12 months to collecting measures of individual levels of local ecological knowledge. In between the two periods, researchers met to make consensual decisions on the structure and content of the data collection protocols.

We obtained free prior and informed consent of each village and individual participating in this study as well as agreement of the political organizations representing the indigenous groups in which we worked. The research adheres to the Code of Ethics of the International Society of Ethnobiology and has received the approval of the ethics committee of the Universitat Autònoma de Barcelona (CEEAH-04102010).

### *Studied Societies*

The three studied societies resemble one another in that (1) they depend on the consumption of local natural resources for their subsistence, generally based on a combination of foraging and farming; and (2) they have recently been integrating with the broader society and monetary economy, although the extent of such changes varies from one society to another. Below, we

Downloaded from fmx.sagepub.com by guest on February 27, 2016

provide some glimpses of the three societies and refer the reader to other published work for additional information.

The Baka in southeastern Cameroon are one of the many hunter-gatherer groups of the Congo Basin. Until the 1950s, they were highly nomadic and depended mainly on wild animals and plants for their livelihoods (Bahuchet et al. 1991), while maintaining economic and social relations with sedentary farming villages. Baka subsistence activities changed after the 1950s, when they began to settle in villages along the roads and to cultivate their own fields (Leclerc 2012). Nowadays, the Baka combine hunting-gathering with farm labor and with cultivation of cassava and plantains, their main staple crops.

The Punan number ~10,000 people living in Indonesian Borneo. Until the 1950s, their traditional economy was based on hunting bearded pigs, preparing starch from hill sago, a wild clump-forming palm, and bartering forest products with locally settled farmers (Kaskija 2012). The Punan started living in more permanent settlements in the mid-1950s, increasingly engaging in wage labor and adopting swidden rice cultivation (Levang et al. 2007). Despite such changes, today the Punan continue to engage in long travel and seasonal stays in the forest for hunting wild boars and gathering forest products such as eaglewood, rattan, and live animals, important sources of cash income (Kaskija 2012; Levang et al. 2007).

The Tsimane' are a hunter-horticulturalist society formed by ~12,000 people living in ~100 villages in Bolivian Amazonia. Until the late 1930s, the Tsimane' maintained a traditional and self-sufficient lifestyle, but their interactions with the Bolivian society have steadily increased since the 1940s (Reyes-García et al. 2005). Previously semi-nomadic, they are now mostly settled in permanent villages. The Tsimane' rely on slash-and-burn farming supplemented by hunting, fishing, gathering, and wage labor in logging camps, cattle ranches, and in the homestead of colonist farmers. Their main cash crops are rice and maize, although the barter of thatch palm also provides an important source of cash income.

## *Development of Knowledge Test*

Researchers invested the first five months of fieldwork in learning the local languages, getting adapted to the local mores, building up trust with participants, collecting background information, and developing and pilot testing the methods to be used. We also collected background information on the content of two domains of knowledge: medicinal plants and hunting. In each site, we conducted 20 free listings on medicinal plants and 20 on game. We also collected semistructured interviews to have a deeper understanding of the

Downloaded from fmx.sagepub.com by guest on February 27, 2016

meaning, values, and beliefs of the studied domains of knowledge. During these interviews, we asked about the most common illnesses and remedies, the behavior of different animals, and the hunting techniques used among others.

Ethnographic information informed the design of the knowledge tests. Since we worked in three culturally and ecologically different contexts, we had to construct site-specific knowledge tests. However, to allow for the cross-cultural comparability of data, we followed the same protocol to generate questions and to structure data collection tools. All of the tools were pilot tested and refined in villages with the same cultural background as the study villages.

## *Methods to Measure Individual Knowledge*

We measured individual levels of medicinal plants and hunting knowledge using four different methods: identification task, structured questionnaire, self-reported skills questionnaire, and peer ratings (Table 1). Data were collected among all adults ($\geq$16 years old) living in two Baka, two Punan, and two Tsimane' villages. Systematic data collection spanned 12 months, a period during which researchers visited each informant several times. The protocols can be accessed at http://icta.uab.cat/etnoecologia/lek.

*Identification Tasks.* Broadly, the identification tasks consisted of asking informants to identify stimuli corresponding to several species. We used free listing results to selected species cited by two or more informants and divided them into three groups according to saliency (frequent, common, and rare). We then randomly chose five items from each group. After testing, the list was reduced to 10 items. In the identification task for medicinal plants, assistants read informants the name of the 10 selected plants and asked them whether they knew the plant, and, if so, whether it had a medicinal use. We created a knowledge score corresponding to the number of plants with medicinal use reported by the informant. In the identification task to assess hunting knowledge, we presented informants with stimuli from a known origin (i.e., a skull provided by the prey's hunter) and asked each informant to provide the vernacular name of the species. The stimuli included pictures, recordings (i.e., a bird's song), and animal parts (i.e., a skull, a feather). Since the stimuli were from a known origin, we generated the hunting scores by contrasting informant's responses with information from the known origin.

*Structured questionnaire.* Based on information collected in the literature from the area and during the ethnographic phase, we developed a set of

Downloaded from fmx.sagepub.com by guest on February 27, 2016

**Table 1.** Methods to Measure Local Ecological Knowledge, by Domain of Knowledge.

| Domain | Task | Definition | Range | Variable name |
|---|---|---|---|---|
| Medicinal plants, $N = 333$ | Identification | Number of medicinal plants recognized (from a list of 10) | 0–10 | MP_Identification |
| | Structured questionnaire | Agreement with the group regarding ailments that can be cured with the 10 medicinal plants | 0–10 | MP_Agreement |
| | Self-reported skills | Index that accounts for the total number of medicinal uses known (out of 10 selected plants) and the last time those were used | 0–20 | MP_Skills |
| | Peer rating | Average rating provided by six evaluators on subject's medicinal knowledge | 0–4 | MP_Rate |
| Hunting, $N = 371$ | Identification | Number of game stimuli recognized (from a list of 10) | 0–10 | Hunt_Identification |
| | Structured questionnaire | Agreement with the group regarding animal behavior/ ecology of 10 game species | 0–10 | Hunt_Agreement |
| | Self-reported skills | Ability to put hunting knowledge into practice | 0–10 | Hunt_Skills |
| | Peer rating | Average rating provided by six evaluators on subject's hunting knowledge | 0–4 | Hunt_Rate |

Downloaded from fmx.sagepub.com by guest on February 27, 2016

questions specific to each domain of knowledge and site. The questions to assess medicinal knowledge referred to the use and preparation of medicinal plants and the questions to assess hunting knowledge referred to the behavioral ecology of the 10 species appearing on the identification task. For example, we asked, "When is the mating season of . . . ?" "Is *telau* more active during the day or the night?" Originally, we designed the questions as multiple choice, but after unsuccessful attempts to apply such questionnaires, we decided to collect data with open responses, which were later recoded into categories. As such, data are not suitable for cultural consensus analysis, we generated a measure of agreement with the group based on the number of times the informant's answer matched the modal response to a question (D'Andrade 1987). As both the medicinal plant and the hunting knowledge tests had 10 questions each, our scores rank from 0 (not a single match with the modal response) to 10 (correct match in all 10 questions).

*Self-reported skills questionnaire.* To measure individual skills, we asked informants to self-report their ability to perform some practices that, according to our ethnographic information, embody local knowledge. For example, to measure skills regarding medicinal plants, we asked informants to report the last time they had prepared the plant remedies they listed in the previous exercise. We created a score on skills using medicinal plants that accounts for the total number of medicinal uses reported by the informant (from the 10 selected plants) and the last time they were used. To assess hunting skills, we asked informants to self-report on hunting frequency, weapons used, and success with difficult-to-catch preys (i.e., sun bear for Punan, tapir for Tsimane', and wild boar for Baka). The hunting skills score was created by assigning points according to such self-reported skills.

*Peer-rating exercise.* To assess how peers evaluate the local knowledge of a given subject, we developed a protocol that aimed to minimize the potential limitations of such a method. First, we selected the people who would conduct the ratings. To minimize biases generated by the evaluator–subject relation, we diversified the group of people evaluating each subject. We grouped households in each studied village into affinity groups (broadly based on kinship and geographic proximity). We then selected one or two household heads in each of those affinity groups to form groups of six evaluators. The selection was done such that each group contained three men and three women and a wide representation of ages. Second, we grouped the names of adults in the sample in lists containing 20 names randomly chosen.

Downloaded from fmx.sagepub.com by guest on February 27, 2016

Each group of evaluators was then assigned one of the lists. In private interviews, we asked each evaluator to rate subjects on the list based on the subject's knowledge regarding medicinal plants and hunting. We first asked, "Who are the best [healers/hunters] in this village?" We assigned four points to the people listed. Then, we read the name of the first subject in the list and asked the evaluator the following questions (adapted to each society): (1) Does [name] know how to cure with plants? and (2) Is [name] a good hunter? Evaluators could rate the person's ability as excellent (4 points), good (3 points), average (2 points), not so specialized (1 point), or does not practice (0 points). If one or more evaluators did not give an evaluation for a given subject, we looked for an additional evaluator, the goal being that ultimately each subject underwent six evaluations. The final knowledge corresponds to the average rate provided by all the evaluators rating the knowledge of a subject in a given domain.

Unfortunately, we could not always keep the protocol unchanged. For cultural reasons, it was difficult to obtain individual evaluations for women's hunting knowledge among the Baka and the Punan, as evaluators simply replied that women do not (know how to) hunt. Thus, we assigned the value of 0 to women in the sample of those societies (0 meaning that the person does not practice the activity).

## *Data Analysis*

We ran a series of pair-wise Pearson correlations of the scores derived from peer ratings against the scores obtained through the other methods. Correlations were done by domain of knowledge and studied society. When using Pearson correlations, we estimate correlation coefficients with Sidak corrections for multiple comparison fallacies, thus making estimates more conservative. As we were unable to obtain all the measures for all subjects, we only report results for those individuals with complete data.

We conducted additional analysis to test whether the magnitude and significance of the correlation coefficient is affected when using subsamples of informants who were rated by (1) three or less informants; (2) more than three; (3) more than four; or (4) more than five. The sample used to test (2) included the subsamples used to test (3) and (4).

To test whether the results vary at different levels of knowledge, we divided the sample in the three following subsamples: (1) informants whose average rating was higher than 0; (2) the 50% of informants with lower ratings; and (3) the 50% of informants with higher ratings. Finally, as the two selected domains of knowledge might reflect gendered activities (men

Downloaded from fmx.sagepub.com by guest on February 27, 2016

**Table 2.** Descriptive Statistics of the Average Rating, per Domain of Knowledge and Society.

| | | *N* | Avg. | *SD* | Min. | Max. |
|---|---|---|---|---|---|---|
| Medicinal plants | Pooled | 333 | 1.27 | 1.05 | 0 | 4 |
| | Baka | 112 | 1.66 | 1.04 | 0 | 4 |
| | Punan | 107 | 0.68 | 0.63 | 0 | 2.6 |
| | Tsimane' | 114 | 1.44 | 1.13 | 0 | 4 |
| Hunting | Pooled | 371 | 1.13 | 1.28 | 0 | 4 |
| | Baka | 155 | 1.04 | 1.30 | 0 | 4 |
| | Punan | 107 | 1.17 | 1.24 | 0 | 4 |
| | Tsimane' | 109 | 1.23 | 1.28 | 0 | 3.8 |

**Table 3.** Results from Pearson Correlations between Peer Ratings of a Subject's Medicinal Plant Knowledge and Other Measures of Medicinal Plant Knowledge.

| | A | B | C | D |
|---|---|---|---|---|
| | Pooled ($N = 333$) | Baka ($N = 112$) | Punan ($N = 107$) | Tsimane' ($N = 114$) |
| MP_Identification | .278** | .404** | .279* | .480** |
| MP_Agreement | −.025 | .347** | .208 | .237 |
| MP_Skills | .438** | .331** | .340** | .379** |

* and **Significant at the .05 and .01 levels.

hunt more and women might gather more plants for medicine) and not knowledge of expertise, we also test whether the magnitude and significance of the correlation coefficient changes for the subsamples of women and men.

## Results

From a potential range between 0 (the person was consistently rated as not performing the activity) to 4 (consistently rated as excellent), average ratings of medicinal plants and hunting knowledge are below the midpoint of 2 (Table 2). In general, the largest variation, as indicated by the magnitude of the standard deviations, was in hunting knowledge ratings.

For the pooled sample, we find a positive and significant correlation between an individual's average rating and two of the three scores of medicinal plant knowledge (Table 3, column A). While modest, the largest

Downloaded from fmx.sagepub.com by guest on February 27, 2016

**Table 4.** Results from Pearson Correlations between Peer Ratings of a Subject's Hunting Knowledge and Other Measures of Hunting Knowledge.

| | A | B | C | D |
|---|---|---|---|---|
| | Pooled ($N = 371$) | Baka ($N = 155$) | Punan ($N = 107$) | Tsimane' ($N = 109$) |
| Hunt_Identification | .353** | .483** | .502** | .393** |
| Hunt_Agreement | .191** | .226* | .196* | .142 |
| Hunt_Skills | .604** | .565** | .744** | .754** |

* and **Significant at the .05 and .01 levels.

correlation coefficient is with the score of the self-reported skills questionnaire (coeff. = .438, $p < .001$). Our measure of agreement with the group was not statistically related to average ratings. For the sample of the three case studies, the individual's average rating correlates with the scores in the identification task and in the self-reported skills questionnaire. Only among the Baka did we find a statistically significant association between agreement with the group and average ratings (Table 3, column B).

Data from hunting knowledge follow a similar tendency. As for medicinal plants knowledge, the largest correlation coefficient was found between average rating and self-reported hunting skills (coeff. = .604, $p < .001$; Table 4, column A). The scores of the identification task and agreement with the group also correlate with average ratings, albeit coefficients are low ($r < .4$, $p < .001$). When considering the group's data, we consistently found that the individual's average ratings correlate with the scores of the self-reported skills questionnaire ($r > .7$ for the Punan and the Tsimane') and with scores from the identification task ($r$ around or above .4). The correlation between average rating and agreement with the group was low and not significant for the Tsimane'.

In Table 5, we test whether the number of evaluators providing ratings affects previous results. Our general finding is that correlation coefficients increased with the number of informants providing ratings, except for the score of agreement with the group, for which we did not find a statistically significant association. In the case of medicinal plants, the coefficient of the association between average ratings and score in the identification task goes from .231 when using the sample of people who were rated by more than three informants ($N = 172$) to .338 when using information for people who were rated by six peers ($N = 90$), despite the reduction in the sample size. An anomaly in this trend is the correlation for the sample of people who

Downloaded from fmx.sagepub.com by guest on February 27, 2016

**Table 5.** Results of Pearson Correlations between Peer Ratings and Other Measures of Medicinal Plants and Hunting Knowledge by Number of Evaluators Providing Ratings.

| | | ≤3 ratings | >3 ratings | >4 ratings | >5 ratings |
|---|---|---|---|---|---|
| Medicinal | | $N = 161$ | $N = 172$ | $N = 110$ | $N = 90$ |
| plants | MP_Identification | .342** | .231* | .297** | .338** |
| knowledge | MP_Agreement | .177 | –.179 | –.093 | –.052 |
| | MP_Skills | .270** | .560** | .637** | .659** |
| Hunting | | $N = 154$ | $N = 217$ | $N = 191$ | $N = 179$ |
| knowledge | Hunt_Identification | –.092 | .416** | .464** | .454** |
| | Hunt_Agreement | .091 | .034 | .013 | .049 |
| | Hunt_Skills | .136 | .562** | .601** | .616** |

* and **Significant at the .05 and .01 levels.

were ranked by three or less people. The trend is more evident when looking at the change in the coefficients with the score of self-reported skills (that range from .270 to .659).

We find a similar trend in hunting knowledge. The correlation coefficients between average ratings and scores in the identification task and the self-reported skills questionnaire were larger when using the sample of people who were rated by six informants ($N = 179$) than when using samples of people rated by fewer informants. Notably, if the subject was rated by less than four informants, the association becomes statistically insignificant.

In Table 6, we present results testing the strength of the association at different levels of knowledge. For medicinal plant knowledge, results from the subsample of informants who were not assigned 0 in the rating resemble overall results (Table 3). When splitting the subsample into two groups, those with the highest and lowest ratings, associations were only statistically significant for the score in the self-reported skills questionnaire, although—for both subsamples—the coefficient is lower than for the average sample. For hunting knowledge, we find correlations of statistical significance between average ratings and the identification tasks and the self-reported skills questionnaire for the subsample of people who were not assigned 0 in the rating exercise and the subsample of people with highest ratings. The coefficients, however, are lower than for the full sample. We find no significant correlation when using the subsample of people with lowest ratings.

Downloaded from fmx.sagepub.com by guest on February 27, 2016

**Table 6.** Results of Pearson Correlations between Peer Ratings and Other Measures of Medicinal Plants and Hunting Knowledge by Level of Knowledge.

| | | Avg. rating higher than 0 | Avg. rating of the lowest 50% | Avg. rating of the highest 50% |
|---|---|---|---|---|
| Medicinal plants knowledge | | $N = 290$ | $N = 183$ | $N = 150$ |
| | MP_Identification | .254** | .095 | .131 |
| | MP_Agreement | –.044 | –.026 | –.082 |
| | MP_Skills | .448** | .213* | .242* |
| Hunting knowledge | | $N = 203$ | $N = 189$ | $N = 182$ |
| | Hunt_Identification | .351** | –.120 | .224* |
| | Hunt_Agreement | –.002 | .123 | –.014 |
| | Hunt_Skills | .500** | .075 | .336** |

* and **Significant at the .05 and .01 levels.

**Table 7.** Results of Pearson Correlations between Peer Ratings and Other Measures of Medicinal Plants and Hunting Knowledge by Sex of Informant.

| | | Female | Male |
|---|---|---|---|
| Medicinal plants knowledge | | $N = 171$ | $N = 162$ |
| | MP_Identification | .304** | .248** |
| | MP_Agreement | .033 | –.093 |
| | MP_Skills | .499** | .377** |
| Hunting knowledge | | $N = 177$ | $N = 173$ |
| | Hunt_Identification | –.026 | .224** |
| | Hunt_Agreement | .311** | .107 |
| | Hunt_Skills | .017 | .243** |

* and **Significant at the .05 and 0.01 levels.

In Table 7, we present results disaggregated by the sex of the respondent. For medicinal plant knowledge, results resemble results from the pooled data, although we find higher coefficients for the subsample of women than for the subsample of men. In contrast, for hunting knowledge, we found different patterns of association between the pooled sample and the subsamples. For the women's subsample, only the measure of agreement with the group was associated with average ratings. Contrarily, for the men's subsample, the correlation was statistically significant for the scores of the identification task and the self-reported skill questionnaire.

Downloaded from fmx.sagepub.com by guest on February 27, 2016

## Discussion

We organize the discussion around three main findings and then present lessons learned during the application of the methodology in the field.

First, we found that peer evaluations seem to correlate with several, but not all, standard measures of local knowledge. Across three societies and two domains of knowledge, peer evaluations correlate with scores in identification tasks and—especially—self-reported skills. Correlations were less significant with scores of agreement with the group. In a seminal paper, Boster and Johnson (1989) highlighted that measures of group agreement might fail to identify experts, as expertise may involve knowledge that is beyond normative knowledge. When calculating agreement with modal responses, the knowledge of an expert hunter might look marginal since responses by the larger number of nonexperts drives what is considered normative. Furthermore, differently than for structured interviews, where everyone gets to fairly respond to the same stimuli, the use of open-ended elicitation does not allow for a fair comparison of individual answers, a flag that might also help explain the lack of association between our measure of agreement with the group and other knowledge measures.

Second, the higher the number of evaluators the higher the correlation coefficient. Since we took special care in selecting groups with evaluators of different age, sex, and from different kinship groups, the finding might indicate that increasing the number of evaluators helps reduce inherent "noise" in judgments. The finding has repercussions for researchers using any kind of peer recommendation (*sensu*, Davis and Wagner 2003) to select informants, as it implies that accurate selection should be based on several recommendations. The tendency in our data is to obtain larger correlation coefficients with each additional evaluator, but since we only interviewed a maximum of six raters we cannot assess which is the ideal number after which each additional evaluator ceases to improve previous estimates or whether such a number is contextual to the domain of knowledge. Future studies might pursue this line of thought in two different ways. First, researchers could obtain ratings for more subjects and draw accumulation curves to assess the optimal number of evaluators for a specific context. Second, researchers could use a network perspective, where experts are those recognized by a wide range of actors who are themselves not connected to one another.

Third, we find higher correlation coefficients when using the full sample than when dividing it into the subsamples with the highest and lowest rating scores. Such results suggest that peer evaluations are more useful in

Downloaded from fmx.sagepub.com by guest on February 27, 2016

differentiating between those with large knowledge differences than in providing fine-grain ranked evaluations. Furthermore, the association was not significant for the subsample with lowest ratings, potentially because of evaluators' underestimations. For example, when Baka men were asked about women's hunting abilities, they often said, "She does not know how to hunt," so we assigned a score of 0. However, when asked directly, the same woman reported occasional hunts. Notice that this explanation fits well with our findings for the subsamples of men and women: Women's rating as hunters, a prominent men's domain in the three groups, does not correlate with their scores in identification tasks and self-reported skills. Differently, women's ratings in medicinal plants, a domain of knowledge in which women might have more expertise, result in higher correlations with the other measures of knowledge. So, cultural stereotypes about, for example, gender division of labor, can lead to underestimations of some individuals' abilities.

We now highlight two methodological lessons learned during the method implementation. First, in the testing phase, we realized that peer evaluations do not work equally well across all domains of knowledge. Furthermore, for some domains of knowledge, it did not work at all. Originally, we had planned to collect ratings on wild edibles and agricultural knowledge. However, we were not able to collect rating data regarding wild edibles in any of the studied societies. Some evaluators considered that the domain of knowledge was too wide to allow for an accurate rating: A person's evaluation in bringing honey would be different from the same subject's evaluation in bringing wild edible fruits. Other evaluators told us that the collection of wild edibles was easy and everybody knew the same things. Regarding agriculture, among both the Baka and the Punan, we were told that it was difficult to evaluate an individual's agricultural knowledge, as agricultural fields were jointly managed by households. While reasons why peer evaluation might not work in a domain of knowledge might vary from one case to another, the overall implication is that researchers need to first find out whether the domain of knowledge is culturally recognized and suitable for ratings. This, of course, can be only done through actual testing in the field.

The second lesson learned relates to drawing people's judgments from peers. Previous research, and our own ethnographic information, suggests that people in the studied societies might evaluate their peers. For example, hunting abilities of young men are carefully scrutinized by young women and their families when choosing a mate (Pillsworth 2008). However, whether it is culturally correct to publicly express such evaluations

Downloaded from fmx.sagepub.com by guest on February 27, 2016

depends on the society, the domain of knowledge, or the relation between the evaluator and the subject among other things. So, in each specific case, researchers should check whether it is culturally appropriate to ask about evaluations and build local trust with people before using peer evaluations.

## Conclusion

Results from this work suggest that peer evaluation can be a reliable measure of individual local knowledge, provided that researchers pay special attention to select a culturally recognized and suitable domain of local knowledge and that they use a large and diverse (in age, sex, and kinship) group of evaluators. Major disadvantages of using peer evaluation as a proxy of individual levels of knowledge are that (1) the results not allow for the study of the domain of knowledge in itself (i.e., we might know who the healer is without knowing the local pharmacopoeia); and (2) the method is not very accurate in providing a fine-grained ranked evaluation of subjects. Conversely, peer evaluation allows researchers to obtain individual level measures of knowledge without digging into the local knowledge of the group. In that sense, peer evaluations can provide a more affordable method in terms of difficulty, time, and budget to obtaining measures that allow for the study of levels of intracultural variation of knowledge.

### Acknowledgments

We extend our deepest gratitude to the Baka, the Punan, and the Tsimane' for their friendship, hospitality, and collaboration. Reyes-García thanks the Dryland Cereals Research Group at ICRISAT-Patancheru for providing office facilities.

### Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement number *FP7-261971-LEK* to Reyes-García.

Downloaded from fmx.sagepub.com by guest on February 27, 2016

## References

Anadón, J. D., A. Giménez, R. Ballestar, and I. Perez. 2009. Evaluation of local ecological knowledge as a method for collecting extensive animal abundance data. *Conservation Biology* 23:617–25.

Atran, S., D. Medin, N. Ross, E. Lynch, V. Vapnarsky, E. U. Ek, J. Coley, C. Timura, and M. Baran. 2002. Folkecology, cultural epidemiology, and the spirit of the commons. A garden experiment in the Maya lowlands, 1991–2001. *Current Anthropology* 43:421–51.

Bahuchet, S., D. McKey, and I. de Garine. 1991. Wild yam revisited: Is independence from agriculture possible for rain forest hunter-gatherers? *Human Ecology* 19:213–43.

Begossi, A. 1996. Use of ecological methods in ethnobotany: Diversity indices. *Economic Botany* 50:280–89.

Berkes, F. 1999. *Sacred Ecology: Traditional Ecological Knowledge and Resource Management*. Philadelphia: Taylor & Francis.

Berkes, F., J. Colding, and C. Folke. 2000. Rediscovery of traditional ecological knowledge as adaptive management. *Ecological Applications* 10:1251–62.

Boster, J. S. 1986. Exchange of varieties and information between Aguaruna manioc cultivators. *American Anthropologist* 88:429–36.

Boster, J. S., and J. Johnson. 1989. Form or function: A comparison of expert and novice judgments of similarity among fish. *American Anthropologist* 91:866–89.

D'Andrade, R. G. 1987. Modal responses and cultural expertise. *American Behavioral Scientist* 31:194–202.

Davis, A., and J. R. Wagner. 2003. Who knows? On the importance of identifying "experts" when researching local ecological knowledge. *Human Ecology* 31:463–89.

Demps, K., F. Zorondo-Rodriguez, C. García, and V. Reyes-García. 2012. Social learning across the lifecycle: Cultural knowledge acquisition for honey hunting among the Jenu Kuruba, India. *Evolution and Human Behavior* 33:460–70.

Gomez-Baggethun, E., S. Mingorria, V. Reyes-García, L. Calvet, and C. Montes. 2010. Traditional ecological knowledge trends in the transition to a market economy: Empirical study in the Doñana natural areas. *Conservation Biology* 24:721–29.

Hewlett, B. S., H. N. Fouts, A. H. Boyette, and B. L. Hewlett. 2011. Social learning among Congo Basin hunter-gatherers. *Philosophical Transactions of the Royal Society of London. Series B* 366:1168–78.

Huntington, H. P. 2000. Using traditional ecological knowledge in science: Methods and applications. *Ecological Applications* 10:1270–74.

Kaskija, L. 2012. *Images of a Forest People, Punan Malinau—Identity, Sociality, and Encapsulation in Borneo*. Uppsala, Sweden: Uppsala University.

Kightley, E., V. Reyes-García, K. Demps, R. Magtanong, V. Ramenzoni, G. Thampy, M. Gueze, and J. R. Stepp. 2013. An empirical comparison of

Downloaded from fmx.sagepub.com by guest on February 27, 2016

knowledge and skill in the context of traditional ecological knowledge. *Journal of Ethnobiology and Ethnomedicine* 9:71. doi: 10.1186/1746-4269-9-71.

Leclerc, C. 2012. *L'Adoption de l'Agriculture chez les Pygmées Baka du Cameroun; Dynamique Sociale et Continuité Structurale*. Paris: FMSH édition.

Levang, P., S. Sitorus, and E. Dounias. 2007. City life in the midst of the forest: A Punan hunter-gatherers' vision of conservation and development. *Ecology and Society* 12:18.

McDade, T. W., V. Reyes-García, P. Blackinton, S. Tanner, T. Huanca, and W. R. Leonard. 2007. Ethnobotanical knowledge is associated with indices of child health in the Bolivian Amazon. *Proceedings of the National Academy of Sciences of the United States of America* 104:6134–39.

Pillsworth, E. G. 2008. Mate preferences among the Shuar of Ecuador: Trait rankings and peer evaluations. *Evolution and Human Behavior* 29:256–67.

Posey, D. 1999. *Cultural and Spiritual Values of Biodiversity*. London: Intermediate Technology.

Reyes-García, V., E. Byron, V. Vadez, R. Godoy, L. Apaza, E. Pérez, W. Leonard, and D. Wilkie. 2004. Measuring culture as shared knowledge: Do data collection formats matter? Cultural knowledge of plant uses among Tsimane' Amerindians, Bolivia. *Field Methods* 16:135–56.

Reyes-García, V., N. Martí-Sanz, T. McDade, S. N. Tanner, and V. Vadez. 2007. Concepts and methods in studies measuring individual ethnobotanical knowledge. *Journal of Ethnobiology* 27:182–203.

Reyes-García, V., V. Vadez, E. Byron, L. Apaza, W. R. Leonard, E. Perez, and D. Wilkie. 2005. Market economy and the loss of folk knowledge of plant uses: Estimates from the Tsimane' of the Bolivian Amazon. *Current Anthropology* 46: 651–56.

Reyes-García, V., V. Vadez, T. Huanca, W. R. Leonard, and T. McDade. 2007. Economic development and local ecological knowledge: A deadlock? Quantitative research from a native Amazonian society. *Human Ecology* 35:371–77.

Romney, A. K., S. Weller, and W. Batchelder. 1986. Culture as consensus: A theory of culture and informant accuracy. *American Anthropologist* 88:313–38.

Ticktin, T., and T. Johns. 2002. Chinanteco management of Aechmea magdalenae: Implications for the use of TEK and TRM in management plans. *Economic Botany* 56:177–91.

Zarger, R., and J. R. Stepp. 2004. Persistence of botanical knowledge among Tzeltal Maya children. *Current Anthropology* 45:413–18.

Zent, S., and L. Maffi. 2010. *Methodology for Developing a Vitality Index of Traditional Environmental Knowledge (VITEK) for the Project "Global Indicators of the Status and Trends of Linguistic Diversity and Traditional Knowledge."* Salt Spring Island, Canada: Terralingua.

All in-text references underlined in blue are linked to publications on ResearchGate, letting you ac

Downloaded from fmx.sagepub.com by guest on February 27, 2016