Culture and Social Norms: Development and Application of a Model for Culturally Contextualized Communication Measurement (MC3M)

Rain W. Liu1*, Maria K. Lapinski2, John M. Kerr2, Jinhua Zhao3, Tsering Bum4 and Zhi Lu5

1University of Arizona, Tucson, AZ, United States, 2Michigan State University, East Lansing, MI, United States, 3Cornell University, Ithaca, NY, United States, 4Emory University, Atlanta, GA, United States, 5Peking University, Beijing, China

Studies of social norms are common in the communication literature and are increasingly focused on cultural dynamics: studying co-cultural groups within national boundaries or comparing countries. Based on the review of the status quo in cross-cultural measurement development and our years of experience in conducting this research among a co-cultural group, this paper describes a Model for Culturally Contextualized Communication Measurement (MC3M) for intercultural and/or cross-cultural communication research. As an exemplar, we report on a program of research applying the model to develop a culturally derived measurement of social norms and the factors impacting the norm-behavior relationship for members of a unique population group (i.e., ethnically Tibetan pastoralists in Western China). The results provide preliminary evidence for the construct validity and reliability of the culturally derived measurements. The implications, benefits, and shortcomings of the MC3M model are discussed. Recommendations for advancing both conceptual and measurement refinement in intercultural and cross-cultural communication research are provided.

Keywords: unique populations, social norms, cross-cultural communication, culturally derived measures, measurement validation

INTRODUCTION

Culture and Social Norms: Development and Application of a Model for Culturally Contextualized Communication Measurement (MC3M)

Social norms research has rapidly garnered popularity in the past several decades in multiple disciplines, such as communication, social psychology, public health, and economics (Chung and Rimal, 2016). Given the power of normative influence on perceptions and actions consistently shown in the body of literature (Borsari and Carey, 2003; Rhodes et al., 2020), social norm theories, rooted in the U.S.-based research, are being applied in numerous cross-cultural contexts (Mackie et al., 2015). Yet, problems persist with inconsistencies in the conceptual and operational definitions of norms (Shulman et al., 2017), and findings of prior studies may be culturally bound (Chung and Rimal, 2016).
To help fill this gap and advance scholarship on social norms and other culturally contextualized communication measurements, we combine qualitative and quantitative approaches to develop culturally derived measures of social norms in a unique population. Specifically, we describe research that examines the nature of interpersonal communication as a basis for shaping social norms and normative perceptions (Lapinski et al., 2021), using this information to derive a series of measures rooted in the cultural context. Data obtained from a multi-year project, including field visits, in-depth interviews, and household surveys about grassland conservation behaviors with ethnically Tibetan pastoralists on the Tibetan Plateau in Western China, are presented as the basis for conceptualization, scale development, and initial evidence for the validity of measures of social norms and related constructs from theories on the communication of social norms (Rimal and Real, 2005; Lapinski et al., 2018). This population is the focus of our work because of their critical role in ecosystem conservation issues in Asia and as a marginalized cultural group (Bessho, 2015; Bum, 2016). As the basis for this research, we offer a model derived from existing research and practice for the development or adaptation of constructs and measures for intercultural or cross-cultural communication research. The value of this paper is the presentation of a model for developing measures of social norms (and related scales in communication studies) accounting for cultural dynamics. This process is useful beyond the particular population studied here, as the detailed steps described in the model shed light on future research on similar issues (e.g., conservation and health) among marginalized groups or populations with unique historical and/or cultural backgrounds (e.g., indigenous people and ethnic minorities).

### Conceptualizing and Measuring Social Norms in Cultural Context

Generally, social norms are “rules and standards that are understood by members of a group, and that guide or constrain social behavior without the force of law” (Cialdini and Trost, 1998, p. 152) shared through interpersonal and mediated communication (Kincaid, 2004). Social norms can influence health, environmental, and philanthropic attitudes and behaviors and can be influenced through communication campaigns (Shulman et al., 2017). International attention has focused on the use of social norm campaigns as key to social change on various issues (e.g., child marriage, female genital mutilation/cutting, food waste, vaccination) because these efforts involve changing beliefs and actions of an entire community or cultural group rather than those of individuals (UNICEF, 2010).

Despite the growing popularity of social norms research, critical issues remain in literature, including vague conceptualizations of what constitutes a social norm and conflated definitions and inadequacies in the measures of different types of norms (Shulman et al., 2017). These problems “ impair our ability to understand what norms are, how they work, how they should be measured, and boundary conditions that dictate where norms should and should not be applied” (Shulman et al., 2017, p.1209). Meanwhile, the increasing trend of social norms research conducted as comparative studies or in countries other than the U.S. and Europe in recent years (e.g., Geber et al., 2019; Stankou et al., 2019) has created a demand for new methods conceptualizing and measuring social norms and related constructs.

Indeed, what we know about norms may be impacted by the so-called WEIRD (Western, Educated, Industrialized, Rich, Democratic) phenomenon documented in psychological research (Henrich et al., 2010). Shulman et al.’s (2017) examination of 832 empirical studies in English language journals found that most studies of social norms (82.4%) were conducted in the U.S. and Western Europe; similar findings exist in global development where few international studies address measurement development or fundamental conceptualization of norms (Mackie et al., 2015).

Constructing valid and reliable measures of key study concepts is regarded as one of the most critical steps in empirical research. No matter how well-designed a study is, poor measurement of study constructs can yield errors in interpreting the results. When studies are designed to compare two cultures or to study communication patterns and processes in a unique population or co-cultural group within a larger group, the measurement challenges are compounded (Croucher and Kelly, 2019, 2020). Differences in the conceptualization of core study ideas, languages, values, and other factors lead to substantial challenges when researchers try to maximize conceptual and measurement equivalence, reliability, and construct validity of measurement for samples from co-cultural groups within national boundaries or across national boundaries (Herdman et al., 1997; Steenkamp and Baumgartner, 1998; Davidov et al., 2018).

Because of the culturally bound nature of social norms, it is crucial for researchers to establish and clearly describe conceptualizations and measurements of norms embedded in the appropriate cultural and social context. By culturally bound, in this case, we mean that although social norms, as unwritten codes of conduct, appear to exist in all human cultures, their form and function vary by group, complicating measurement. A lack of culturally valid measurement may hinder progress in theory building, especially in identifying boundary conditions for theories.

Studies of social norms and cultural dynamics have focused on nation/country (e.g., Cialdini et al., 1999) or race/ethnicity (e.g., LaBrie et al., 2012) as a delimiting concept. We recognize the benefits and limitations of using country or nation as the sole proxy or operationalization of culture, despite the prevalence of this practice in cross-cultural research (c.f., Schaffer and Riordan, 2003).

Using country, race, or ethnicity to identify cultural groups is convenient, clear, and tidy; most people can self-identify these
characteristics when asked with valid indicators in the measures. Yet, country and culture are incongruent under most conditions. Generally, multiple co-cultural groups exist under the same overarching national identity (Orbe, 1997). As such, culture may function at the level of a nation-state, a co-cultural group within a nation-state, or any collective of people who share deep or surface-level cultural elements (termed a unique population). For the current study, we draw from the intercultural communication literature and use the term “culture” to include communities of people with uniquely shared communication characteristics, perceptions, values, beliefs, and practices. Shared practices, ethnicity, and language serve as indicators for the cultural group, which is the focus of the present study; ethnically Tibetan pastoralists. This group shares the following characteristics: they are historically nomadic and engage in animal husbandry, and they have Tibetan ethnicity with the Kham Tibetan dialect as their primary language. Because, fundamentally, culture influences how people view the world, identifying within-culture conceptualizations of key study constructs should be the first step in empirical inquiry. As unwritten implicit rules, social norms are formed, shaped, and reinforced through observation and interpersonal and mediated communication among a collective. Normative perceptions may be formed about both the prevalence of behavior (i.e., commonly called descriptive norms; what is done by most members of a group) and what most people think to be appropriate or inappropriate behaviors (i.e., injunctive norms; what is socially approved or disapproved; Cialdini et al., 1990). Hence, it is critical to acknowledge the socially and culturally shared nature of social norms, as people relate to in-group members within a specific culture. That is, social norms, by their nature, emanate from collectives within a system. As such, it is necessary to identify the influential people and in-groups who are most connected to particular decisions or behaviors in order to contextualize norms.

Some research demonstrates the culturally bound nature of conceptualizations of social norms and their communication (e.g., Jensen and Bute, 2010; Lapinski et al., 2015). Using in-depth interviews and observation, the literature indicates that key conceptualizations developed in one cultural context (like injunctive norms with social prescriptions for appropriate behavior) may not exist in the same form when examined through a different cultural lens (Jensen and Bute, 2010). Likewise, the nature of interpersonal and mediated communication about what is approved behavior is constrained by the nature of the social system (Elwood et al., 2000; Lapinski et al., 2015) and connected to cultural predispositions (Lapinski et al., 2019).

Developing culturally derived social norms measures is also critical to enhance both the internal and external validity of the existing corpus of research to account for culturally-based concepts and processes (Mollen et al., 2010). Surprisingly little is written about how to develop reliable and valid culturally derived measures of communication concepts like social norms; instead, one must go to the literature in cross-cultural and organizational psychology to find scholarship addressing some of these issues (c.f., Schaffer and Riordan, 2003). In public health, there is a robust literature on the cross-cultural adaptation of scales; yet, Epstein et al. (2015) reviewed 31 studies making recommendations for cross-cultural adaptation (CCA) and concluded there was no consensus on best practices for adapting measures across cultural contexts.

In sum, identifying and refining the culturally derived conceptualization of social norms is the first step in developing methods for measuring these constructs. Measurement development is critical for expanding social norms research to account for cultural similarities and differences in order to enhance both internal and external validity in the corpus of research to account for culturally-based concepts and processes (Mollen et al., 2010; Lapinski et al., 2019).

Studies of Social Norms in Cultural Context: Absolutism, Universalism, and Relativism
Various approaches to studying cultural dynamics in social normative influence are evidenced in the literature (c.f., Fischer et al., 2009; Lee and Green, 1991; Park and Levine, 1999). Many of these studies have involved comparative research designs in which data from a U.S. sample are compared to a sample(s) of people from another nation (Shulman et al., 2017). The predominant theories that address social norms, such as the theory of reasoned action (TRA; Fishbein and Ajzen, 1975), the theory of planned behavior (TPB; Ajzen, 1991), focus theory of normative conduct (Cialdini et al., 1990), social norms approach (SNA; Berkowitz, 2004), and theory of normative social behavior (TNSB; Rimal and Real, 2005), have been developed and tested primarily in the U.S. with measures of the core theoretical concepts constructed in English. Studies using these theories sometimes provide evidence for measurement reliability and validity of the study measures using data from samples, often of college undergraduates, in various regions of the U.S. (e.g., Cialdini et al., 1999; Jang, 2012).

It is when these theories and measures are applied in new cultural contexts that challenges may arise. That is, by moving existing normative concepts and measures into new cultural contexts, studies may fail to account for the dynamics of normative influence unique to the new context. A framework in cross-cultural psychology that can be applied to communication research describes three orientations to the cross-cultural adaptation of theories and measures, including absolutism, universalism, and relativism (Herdman et al., 1997; Berry et al., 2002). Based on this framework, there are roughly three approaches to studying social norms in cultural context: 1) adoption of the conceptualization and measures from existing theories and using them with no modification in a new cultural context (absolutism); 2) using conceptualization and measures developed in one cultural context (often in the language of the researcher) and translating the measures into the primary language of the study participants or making other adjustments for cultural context (universalism), and 3) developing the study concepts and measures based on data (or dialogue) from within the
cultural context in the language of participants for each cultural group included in the study (relativism). In each of these cases, the nuances of the study procedures and the reporting of the processes are different for each study. For example, studies may or may not report on: the development of conceptual definitions, translation and back translation of items, evidence for scale reliability or validity, or measurement invariance. In the following, we review and summarize examples of these orientations from across disciplines and then propose a series of recommended practices derived from the existing literature, for culturally derived measurement of communication constructs.

**Absolutism**

**Absolutism orientation** assumes a minimal impact of “culture” on the constructs being studied (i.e., they are culture-free) because of the species-wide similarities among all human beings. As a result, standard instruments measuring the focal constructs are considered appropriate to be used in different cultures. This practice may result in a construct conceptualized and operationalized in one culture that is “imposed” directly onto another culture (Berry et al., 2002). It involves adopting the conceptual definitions, study materials, and measures directly from prior research without substantial modifications. It may include using measures from prior research in a particular country without any translation procedures or evidence for measurement construct validity or equivalence (e.g., Thøgersen and Ölander, 2006; Abikoye and Olley, 2012; Nguyen and Neighbors, 2013; Savani et al., 2015).

For example, Bobek et al. (2007) conducted an experimental study with participants recruited from Australia, Singapore, and the U.S. to examine the effects of social norms on tax compliance using Cialdini and Trost’s (1998) taxonomy of social norms. Factor analysis and scale reliability analysis were performed to establish evidence for the scales’ validity and reliability before proceeding to test hypotheses. However, across the three national samples, the constructs and measures were assumed to be equivalent, and a translation process was not described.

Likewise, using measures from the theory of planned behavior (TPB; Ajzen, 1991), Wan et al. (2018) examined the moderating effect of subjective norms on the behavioral intention of using urban green spaces among Hong Kong residents. The convergent and discriminant validity and reliability of the measures were assessed before testing the structural model. But, no survey translation information was described, although most people in Hong Kong speak Cantonese as their primary language, and only 4.3% of the population use English regularly (GovHK, 2020).

---

1As a caveat, only studies published in English language journals are reviewed here. Further, we only use the information available about these studies in the published version of the paper which may be incomplete.

2Minor adaptations of the scales may be involved to fit with the specific study scenarios or focal behaviors.

3Although English is the dominant language in Australia and the U.S., Singapore’s national languages are English, Mandarin, Malay, and Tamil (Department of Statistics Singapore, 2019).

**Universalism**

The universalism orientation acknowledges that culture substantially impacts how constructs are expressed and defined across cultures. Though this approach still assumes species-wide similarities (i.e., universal patterns), it accepts the idea that measurement needs to be adapted cross-culturally, given that the context-free constructs and measurements are difficult or impossible to obtain. In this approach, conceptual definitions and measures are developed in one cultural context, typically in English. Then the study materials and measures are translated into the country’s language in which the research is conducted. Evidence for back-translation, construct validity, and measurement equivalence may or may not be described. There are a few social norms studies that account for cultural dynamics using this method (e.g., Cialdini et al., 1999; Park and Levine, 1999; Boer and Westhoff, 2006; Fornara et al., 2011; Jang et al., 2013; Stamkou et al., 2019; Walter et al., 2019).

For example, Stamkou et al. (2019) examined the moderating effect of cultural collectivism and tightness on responses to norm violators in 19 countries. The conceptual definition of the key study constructs and the measures, including social norms, norm violations, individualism-collectivism, and tightness-looseness, were adapted from existing literature developed in the U.S and translated into each country’s official language following the procedures outlined by Brislin (1986); validity and reliability evidence was provided. Likewise, Jain et al. (2018) investigated the effect of descriptive and injunctive norms on condom use among young men in Ethiopia using norms measures from the TNSB (Rimal and Real, 2005) translated into Amharic, Afan Oromo, and Tigrigna. Adaptations were made in the norm measures to account for cultural context, but measurement validity and reliability evidence was not presented. Limaye et al. (2012) reported similar process in Malawi; acceptable reliability of the scales was presented, but measurement validity evidence was not included.

**Relativism**

The last orientation, relativism, assumes that because of the substantial role of culture in people’s cognitive thinking patterns and behaviors, it is impossible to use standard measurements across cultures; hence, local instruments developed within a specific culture should be adopted (Herdman et al., 1997; Berry et al., 2002). In this approach, the conceptual definitions and measures are developed within the focal cultural group, often through collaborative processes and formative data collection. The language in which they are developed may be that of the focal country or region. Measurement construct validity and equivalence evidence may or may not be described (e.g., Babalola, 2007; Rimal et al., 2015; Yilma et al., 2020). For example, Rimal et al. (2019) developed a personal narrative-based intervention, including social norms messages targeting adolescent students in Serbia, to improve their driving behaviors using conceptual definitions and measurement based on theory and cultural context.
Formative data (i.e., one-on-one interviews, focus groups, and reaction interviews) was conducted first to develop the intervention and the measures of core concepts, including descriptive and injunctive norms. Results showed acceptable reliability of the normative scales, but measurement validity evidence was not included.

In sum, the literature on social norms and cultural dynamics indicates a range of approaches to developing concepts and measurements in cultural context for both single and multi-culture studies.

Model for Culturally Contextualized Communication Measurement (MC3M)

Based on the research on culturally derived measurement (Hui and Triandis, 1985; Pedhazur and Schmelkin, 2013; Schaffer and Riordan, 2003; Steenkamp and Baumgartner, 1998), research on measurement model validation and equivalence (Bollen, 2005), and our team’s international and cross-cultural research, we present a Model for Culturally Contextualized Communication Measurement (MC3M) containing a series steps for the development of quantitative measures in communication science taking a relativistic approach (Figure 1) and use a variant of this model in the current research. Although we focus here specifically on social norms, we believe this model may benefit other communication research. In the following sections, we describe a series of studies to illustrate the process of applying the model to develop culturally derived social norm measures.

The program of research that we report here was conducted on the Tibetan Platea in the Tsangsum Yangyu (Tibetan) or Sanjiangyuan (Mandarin) area of China, located in southern Qinghai Province. This region is home to about 960,000 inhabitants, 90% of whom are ethnically Tibetan, and nearly 70% are pastoralists, sometimes nomadic, herding mainly yaks and sheep (see Appendix A). Geographically, the territory is vast, with human settlements dispersed, making data collection in the region challenging. The terrain includes glaciers and high-altitude grasslands, which input to three of Asia’s major rivers, the Yellow, Yangtze, and Mekong providing freshwater to nearly a quarter of the world’s population. The population of this region is generally Tibetan Buddhist. Their position as a unique or co-cultural group within China makes Tibetan pastoralists an important group to study social influence processes. They play a key role in the future of this ecologically sensitive region, but studies conducted in this area are rare (Shen and Tan, 2012).

Step 1: Identification of Key Constructs

Discussions with cultural informants, review of the scientific and gray literature about the study region, field visits, and collaborative discussions with project partners were the first stage of this project; Step 1 in the MC3M. The cross-cultural (U.S. multi-ethnic, Han, Tibetan team), cross-disciplinary (anthropology, communication, sustainability, conservation biology, economics) team shared an interest in interpersonal communication about social norms and their effects on conservation behavior and the role of financial incentives in promoting conservation behavior among ethnically Tibetan pastoralists.

The exploratory work conducted in Step 1 revealed results in many key activities and insights, two of which we highlight here. First, discussion with collaborators coupled with our searches of the scientific literature revealed little social science data on the population of interest. This is critical because it drove our approach to the methods we used throughout the remainder of the project. Second, the focal constructs, behaviors, core theory, and research questions/predictions were developed collaboratively based on this process. Animal husbandry behaviors and their impact on the grassland and water ecology were identified as both salient for the study population and conservation practice. Specifically, herding types of animals with less relative ecological impact, reducing herd size to have less impact on grassland quality, and modifying grazing patterns to protect sensitive areas were the behaviors examined; organized patrolling to reduce poaching of wild animals was also examined but is reported elsewhere.

![Figure 1](image-url)
Step 2: In-Depth Interviews

As the next step in developing measures of the normative dimensions and providing construct validity evidence in this cultural context, in-depth interviews were conducted (Step 2 in the MC³M). The purpose of the interviews was to determine whether or not and how normative information was communicated to members of our study population and the character of that information in order to identify conceptualizations of social norms. In addition, we sought to understand the conditions under which normative information was available, the people from whom normative information emanates, and expected outcomes for the focal behaviors. Eighty in-depth interviews were conducted with members of our study population; detailed results are reported in companion papers (Lapinski et al., 2018; Lapinski et al., 2021). Interview data were analyzed via quantitative content analysis, thematic analysis, and network analysis.

The interviews provided the basis for understanding indigenous conceptualizations of injunctive and descriptive norms, outcome expectancies associated with the behaviors, important referent groups for information about our study topics. In brief, the findings from the interviews uncovered normative influence as one basis for social power (Kelman, 1961) among members of the study community (Lapinski et al., 2018) and three essential themes for conceptualizing social norms (Lapinski et al., 2021): 1) a shared understanding of what the participants believe is typical in the community, particularly local herding groups or villages (descriptive norms); 2) what participants believe is approved and disapproved or expected in the community (injunctive norms), and the anticipated reactions of others to compliance or noncompliance with expectations; and 3) important referent groups for decisions about herding (normative referents). Key referents were identified as dependent on the nature of information (general information, advice-seeking, or problem-focused), including herding group members, other villagers, family, and people in positions of power (e.g., veterinarians, government officials, village leaders).

Step 3: Refining Conceptualizations

Based on the findings from the interviews, revised conceptual definitions (Step 3 in the MC³M) and quantitative items were developed (Step 4, described in the method) to investigate further the influences of social norms on behaviors guided by several existing theories of social norms (Fishbein and Ajzen, 1975; Cialdini et al., 1990; Rimal and Real, 2005) and our prior research (Lapinski et al., 2018). Based on the interview data, the conceptualizations of both normative constructs (i.e., perceived descriptive norms and perceived injunctive norms, provided earlier in this paper) have been modified slightly to be culturally appropriate. Consistent with prior research, perceived descriptive norms are conceptualized as pastoralists’ perceptions of the prevalence of referent others’ (herding group and village group member) behavior. Perceived injunctive norms are conceptualized as perceptions of the referent others’ opinions and expectations about behaviors. A common element in conceptualizations of social norms—that social sanctions exist for noncompliance with the norm—was not included in the definition because it was not evidenced in our data. The key referent groups for this behavior are the herding group (if the pastoralist belongs to one) or others from the same village (if the pastoralist does not herd with a herding group). Families have been incorporated into the herding group conceptualization, given the clear overlap revealed from the interview data between these two groups.

Outcome expectations, as well as group identification and group orientation, were considered as key constructs in the study because prior research has shown they enhance the influence of social norms and appear to be critical in studies of cultural dynamics (Cruz et al., 2000; Lapinski et al., 2007) and conceptualizations were shaped based on the results of the in-depth interviews. Outcome expectation is conceptualized as beliefs of the potential losses or benefits related to the behavior and includes monetary and non-monetary outcomes. The types of outcomes identified in the interviews included changes to the grassland, changes to economic well-being, and changes to identity as a Tibetan (Lapinski et al., 2021). Group identity refers to feelings of affinity with one’s social group and the desire to be connected to that group (Rimal and Real, 2005). Group orientation refers to one’s connection to the collective (i.e., the extent to which one’s social groups are central to the decision-making process). Giving priority to group goals over personal goals may function to enhance the influence of social norms on behaviors since group-oriented individuals are guided by group goals and norms in order to maintain harmony within groups (Lapinski et al., 2007). Finally, we conceptualized behavioral intention as a person’s readiness to perform a behavior (Fishbein and Ajzen, 1975) and a possible outcome of normative influence.

These conceptualizations form the basis for the development of items designed to measure each of the constructs. A cross-sectional survey was conducted with our study population in order to complete Steps 4–7 in the MC³M: The hypothesis is proposed:

H: The measures of perceived descriptive norms (PDN), perceived injunctive norms (PIN), outcome expectations (OE), group identity (GID), group orientation (GO), and behavioral intentions (BI) will yield valid and reliable unidimensional scales.

METHODS

Sampling and Participants

Participants were recruited from one city and three counties in the study region via network sampling by project partners (see Appendix A). Yushu Prefecture is an area of 267,000 square kilometers, with a total population of 283,100 people (95.3 percent Tibetan). As of 2015, Yushu Prefecture has one city and five counties; our sample included: Yushu City, Zaduo County, Nangqian County, and Chengduo County. Because of the behaviors examined in this study, three filter questions were asked at the beginning of the survey to ensure that the participant 1) was a pastoralist, 2) with at least 10 yaks in their herd, and 3) was the primary decision-maker in the household (i.e., the head of

Frontiers in Communication | www.frontiersin.org 6 January 2022 | Volume 6 | Article 770513

Liu et al.
the household). Only people who answered affirmatively to these questions were included in the sample. During data cleaning, one participant was removed from the data analysis because his household had fewer than 10 yaks.

In total, 360 Tibetan pastoralists (85% male) in 10 townships participated in the surveys, with an average age of 45.85 (SD = 12.29), ranging from 18 to 80. The average size of the household was 6.52 (SD = 2.57), with an average number of 2.36 (SD = 1.48) school-aged children and 2.31 (SD = 1.48) family members who helped with herding. Regarding the level of education, on average, participants had 1.3 years (SD = 2.36) of schooling (including public schools and monastery schools), ranging from 0 (illiterate; 68.1%) to 15 years. Nearly all (98.3%) reported owning only yaks; less than 1% had both yaks and sheep (three missing responses). The average herd size of yaks was 40.87 (SD = 28.27), ranging from 0 to 200. Approximately 20% of the participants (n = 71) belonged to herding groups, and 9 (12.7%) of them reported themselves as the leader of the herding group.

**Survey Instrument Development**

**Step 4: Initial Item Development and Cognitive Interviews**

The survey items were developed by the project team based on the results of the in-depth interviews (Lapinski et al., 2018, 2021) and prior research on social norms-related variables (Step 4 in the MC3M). The scale items were developed via the procedures suggested by Hunter and Gerbing (1982). Items were developed for each distinct dimension by examining the conceptual definitions of the constructs and by deriving content from the interviews. Multiple items were created for each construct in order to allow for subsequent statistical tests of construct validity (Hunter and Gerbing, 1982). The item construction process resulted in a large pool of items reviewed for face validity by the researchers. To enhance conceptual equivalence (Herdman et al., 1997), each question was discussed by study team members and revised based on the discussion. Items that matched the conceptual definition of the construct were retained. The measures were developed in English and Tibetan simultaneously, captured in English, and then translated into Tibetan with flexibility for local variations in the dialect. The instrument was then back-translated to English to check for accuracy in interpretation and to avoid cultural biases. Then, the study team members discussed the final version of questionnaire questions one by one (see Appendix B for the detailed procedures of translation and back-translation).

Two groups of cognitive interviews (four participants per group) were conducted with local community members to pilot the survey instrument before the data collection. This qualitative approach, conducted prior to the quantitative data collection, helped researchers examine how the respondents process and interpret questions and identify the factors influencing their answers (Cabral and Savageau, 2013). Due to the benefit of improving item interpretation and strengthening scale quality shown in numerous studies (e.g., Collins, 2003; Ryan et al., 2012), the cognitive interview has been recommended as a standard step in survey development, refinement, and adaptation.

During the cognitive interviews, participants were asked to evaluate the survey questions with the goal of increasing the clarity, meaningfulness, and cultural appropriateness of the questions. Modifications were made to question wording and question order, and some questions were eliminated. Although we developed the scales to use verbally administered Likert-type response scales ranging from 1 (strongly disagree) to 5 (strongly agree), based on the suggestions from local collaborators and cognitive interview participants, we adopted the strategy of using fingers (digits; commonly used among people in the sample in everyday life) as a response scale when asking about Likert-type questions (e.g., thumb = strongly agree; the little finger = strongly disagree), to help participant better understand the options. A “Not Sure” option was added based on the suggestions from the local collaborator and the feedback generated from the cognitive interviews.

**Procedures**

Surveys were conducted by four ethnically-Tibetan enumerators who were native speakers of the Kham Tibetan dialect and also fluent in Mandarin Chinese. Enumerators received training on survey skills, survey instruments, and the protection of human subjects by the study team (Step 5 in the MC3M). The enumerators verbally administered all questions using the digit response scale described above and recorded the responses in booklets due to the low level of literacy among our potential participants based on the exciting literature (e.g., John, 2000; Bangsbo, 2008), the fieldwork of our community collaborators in our study area over the years, and data from our previous interviews. To minimize unintended enumerator effects on the survey data, enumerators were trained not to provide any explanations to the survey questions other than clarification or to provide verbal or nonverbal reactions toward participants' answers. Statistical analysis was conducted to ensure that no significant differences existed in study variables for different enumerators.

Upon approaching a potential participant, each enumerator first introduced him/herself and the purpose of the survey briefly. If the individual agreed to answer the initial eligibility questions, the enumerator would record the sex of the respondent through observation first and then ask the three filter questions mentioned above (i.e., a pastoralist with at least 10 yaks who is the head of their household). Once the participant was determined as eligible for the survey, the enumerator proceeded with the informed consent process, adapted to be culturally appropriate while retaining the key elements of consent. Participants were also provided with opportunities to ask questions before deciding to...
TABLE 1 | The Measurement Model of the Six Constructs.

| Construct                           | Indicators                                                                 | Factor loadings | All (α) | HG (α) | VG (α) | CR  |
|-------------------------------------|-----------------------------------------------------------------------------|-----------------|---------|--------|--------|-----|
| **Perceived Descriptive Norms (PDN)** | PDN1- Most people in my herding group/village have reduced the size of their herd over the last 5 years | 0.90            | 0.89    | 0.93   | 0.87   | 0.91 |
|                                     | PDN2- In the last 5 years, people in my herding group/village have reduced the number of yaks they herd | 0.90            |         |        |        |     |
|                                     | PDN3- People in my herding group/village have decreased the number of yaks they herd | 0.86            |         |        |        |     |
|                                     | *PDN4- About how many yaks do most households in your herding group/ village own? | 0.51            |         |        |        |     |
| **Perceived Injunctive Norms (PIN)** | PIN1- I think people in my herding group/village would be upset with me if I don’t reduce the number of animals I graze | 0.93            | 0.84    | 0.79   | 0.83   | 0.93 |
|                                     | PIN2- People in my herding group/village will be dissatisfied with me if I don’t reduce my herd size | 0.92            |         |        |        |     |
|                                     | *PIN3- People in my herding group/village think it is better for me to have a larger herd. (R) | 0.24            |         |        |        |     |
|                                     | *PIN4- People in my herding group/village think it is best for me to reduce my herd size | 0.29            |         |        |        |     |
| **Group Identity (GID)**            | On the whole, how similar do you think you are to most people in your herding group/village – |                |         |        |        |     |
|                                     | GID1- in your values                                                         | 0.83            | 0.82    | 0.82   | 0.81   | 0.86 |
|                                     | GID2- in your behaviors                                                      | 0.89            |         |        |        |     |
|                                     | GID3- in your likes and dislikes                                              | 0.81            |         |        |        |     |
|                                     | *GID4- I feel close to other people in my herding group/village              | 0.45            |         |        |        |     |
| **Group Orientation (GO)**          | GO1- I would do what would please my family                                  | 0.67            | 0.68    | 0.70   | 0.64   | 0.77 |
|                                     | GO2- I usually sacrifice my self-interest for the benefit of others around me | 0.59            |         |        |        |     |
|                                     | GO3- My happiness depends very much on the happiness of those around me       | 0.68            |         |        |        |     |
|                                     | GO4- It is important for me to maintain harmony with people around me         | 0.75            |         |        |        |     |
| **Outcome Expectations**            | *OE1- Reducing the number of yaks I herd will make it hard for me to support my family | 0.28            | NA      | NA     | NA     | NA  |
|                                     | *OE2- Reducing the number of yaks I herd means I wouldn’t have enough yak products (butter, milk) for my family | **              |         |        |        |     |
|                                     | *OE3- If I reduce the number of yaks I herd, I will no longer feel like a “Tibetan.” | 0.40            |         |        |        |     |
|                                     | *OE4- If I reduce my herd size, it will improve the quality of the grassland. (R) | 0.18            |         |        |        |     |
| **Behavioral Intentions**           | *BI1- Over the next 5 years I plan to reduce the number of yaks I herd      | 0.66            | 0.60    | 0.38   | 0.58   | 0.64 |
|                                     | *BI2- I plan to reduce my herd size in the next 5 years                      | 0.76            |         |        |        |     |

R = reverse-coded item; α = Cronbach’s Coefficient Alpha; HG, Herding Group; VG, Village Group; CR, Composite Reliability; *Item dropped when failed tests of internal consistency; ** negative factor loading yielded; NA, not applicable.

participant or not. If they agreed to participate, the enumerator would proceed to the main survey questions. First, each participant was asked if he/she belonged to a herding group. Based on the participant’s answer to this question, he/she was directed to the subsequent questions associated with a specific referent group (people in my herding group vs people in my village), measuring their perceived descriptive norms, perceived injunctive norms, group orientation, group identity, perceived outcome expectation, behavioral intentions of reducing their herd size and demographics. Based on local norms, participants did not receive incentives for participation.

Surveys were conducted in semi-private settings in Kham Tibetan dialect and lasted approximately 30 min each. Participants’ responses to each question were recorded on the survey paper in Mandarin Chinese by the surveyors and manually entered into the computer later by two research assistants who were fluent in both Chinese and English. Each research assistant first entered all the survey data independently, and then their data entry files were carefully compared to identify any inconsistencies caused by human error during the data entry process. Following several days of data collection, data were reviewed, and procedures were discussed to determine whether modifications were necessary; all study procedures were retained. One researcher who was tri-lingual (Kham Tibetan, Mandarin, and English) was responsible for quality control of the procedures and data. All procedures were approved by a university institutional review board.

**Measurement**

Likert-type scales ranging from 1 (strongly disagree) to 5 (strongly agree) were adopted, with an additional option “Not Sure” added; response scales were administered using the enumerators’ fingers as a guide. All survey items (see Appendix D), including factor loadings, are presented in Table 1. Items either focused on herding group members or village group members as the referent, the 5 years prior to the survey as the time period, and herd size reduction as the behavior. Because of the nature of the study procedures, which were...
conducted in the field in naturalistic conditions, without incentives, every effort was made to streamline the questionnaire content and number of items per dimension in order to avoid attrition. For all scales, items retained following confirmatory factor analysis were summed such that higher scores indicated greater levels of the variable.

Establishment of Measurement Model
Based on Hunter and Gerbing (1982), the development and evaluation of a measurement model via factor analysis procedures included three steps: 1) construction of the model, 2) estimation of the observed correlations among the variables/items in the model, and 3) comparison of the observed correlations among variables with the correlations predicted by the model. The measurement model was specified first based on a theory of the relationships among the items. Thus, it was appropriate to use confirmatory factor analysis (CFA) procedures to estimate the parameters of the models and provide construct validity evidence. These procedures are included in Step 6 of the MC3Mi n for this study based on the in-depth interviews.

Scales and Items
Perceived descriptive norms (PDN). Participants’ perceived prevalence of others’ behavior of reducing the herd size among their referent group (herding group or people in the same village) was assessed with four items. One item directly asking about how many yaks they think the most households in their herding group/village own was dropped as it failed the internal consistency test with a low factor loading.

Perceived injunctive norms (PIN). Participants’ perceptions of the referent others’ opinions and anticipations of them reducing the size of their herds were assessed with four items initially. Two items, including a reverse-coded item, were eliminated due to low factor loadings.

Group identity (GID). Participants’ perceived attitudinal similarity and closeness with their referent group (their herding group or people in the same village) was assessed with four items derived from Rimal and Real (2005). One item measuring participants’ perceived closeness to their herding group/village was dropped as it failed the internal consistency test with a low factor loading.

Outcome expectations (OE). Expectations about behavioral outcomes were measured by four items, including a reverse-coded item measuring the perceived benefits associated with herd reduction behavior. The results indicated small correlations among all the items (see Table 1). Hence, it was deemed inappropriate to compose the variable by summing the items. This variable was removed from the rest of the analysis assessing the validity and reliability of the scales.

Group orientation (GO). The extent to which one is oriented toward group goals as opposed to individual goals was measured by a four-item scale derived from Triandis’ (1995) individualism-collectivism (INDCOL) scale and prior research (Lapinski et al., 2007), which has been modified for this study based on the in-depth interviews.

Behavioral intention (BI). Participants’ intent to engage in the study behavior of reducing the number of yaks in their herds was measured with three items initially, including a reverse-coded item measuring the intention to increase the number of yaks in their herds. One item was eliminated due to its low factor loading.

Demographics. Participants’ demographic information was collected at the end of the survey, including biological sex (observed and recorded by the enumerator), age, number of people in their households, number of children, level of education, and residence location (county and township).

RESULTS
Missing Data and “Not-Sure” Responses
Missing data and responses of “not sure” (NS) were scrutinized for patterns (Rubin, 1976) because the population under study is rarely surveyed, and the scales are newly developed (see detailed results in Appendix C). The findings show that NS answers are more prevalent among village groups than herding groups, accounting for 93.62% of the total NS answers, suggesting the influential power of one’s herding group as the source of clearer normative information. For measurement validation in the subsequent analyses, both the missing and the NS data were eliminated, and the pairwise deletion was employed to retain sufficient statistical power.

Construct Validity Assessment
CFA was conducted using the lessR package developed by Gerbing (2021) within R programming environment to provide evidence that the observed scale items measured the same theoretical constructs. Both internal consistency and parallelism (Hunter and Gerbing, 1982) were tested to evaluate the unidimensionality of the measurement model. The a priori specified criteria for item retention for tests of internal consistency include both the pattern and magnitude of the errors between predicted and obtained correlations between items ($e < 0.20$) and examination of the size of the factor loadings. Once items were eliminated from a factor, factors were reanalyzed to test the unidimensionality of the new factor. Behavioral intentions with three items was not included in this test.

In testing the internal consistency among items designed to measure PDN, item #4 was dropped as it failed the internal consistency test with a low factor loading and large error for predicted and obtained inter-item correlations ($e > 0.20$). Since there were only three items left after the elimination, this factor was not tested again for internal consistency. When testing items measuring PIN, items #3 (reverse-coded) and #4 were eliminated.

---

7Based on the suggestions from our experienced local collaborators and cultural insiders, we had to keep the survey short by limiting the number of items for each scale as much as possible, due to the reasons that 1) our survey was verbally administered, which took a much longer time to complete compared to a written/online survey, and 2) our study group had never participated in any studies or completed any surveys. Items developed in each scale with closely shared meaning may confuse them when answering the questions.
due to the low factor loadings and large errors yielded. Two items were retained. Likewise, when testing items measuring group identity, item #4 was eliminated due to the low factor loading and large error. As such, no further internal consistency test was conducted. For the items measuring OE, the results showed insufficient factor loadings of all items developed in this scale with large errors. Hence, we deemed it was inappropriate to compose the variable by summing up the items and removed this variable from the rest of the analysis.

For the items measuring GO, the test of internal consistency via CFA indicated a plausible four-item solution for the scale; all items were retained. All errors for predicted and obtained inter-item correlations were small (e < 0.20, goodness of fit RMSE = 0.06).

Tests of parallelism were next conducted to estimate how items measuring the same factor are distinct from other factors. Instead of assessing macro-level correlations between scales, tests of parallelism are conducted at the level of individual items with a low tolerance for errors (i.e., the discrepancy between the predicted correlations and the observed correlations). Results from the parallelism test showed that the four-factor model solution was acceptable: Comparative Fit Index (CFI) = 0.94, Tucker-Lewis Index (TLI) = 0.91, Root Mean Square Error of Approximation (RMSEA) = 0.07, Standardized Root Mean Square Residual (SRMR) = 0.06, \( \chi^2(67) = 228.01, p < 0.00 \), all errors were below the a priori specified value of 0.20. The factor loading for each scale item was reported in Table 1, in which the five-factor solution was clearly demonstrated.

**Discriminant Validity of the Constructs**

After establishing the measurement model, the relationships among the four constructs were examined to assess the discriminant validity, which refers to measurement items within different constructs that should be unrelated (Hunter and Gerbing, 1982). See Table 2 for the correlations among the variables in both herding and village groups. The mean and standard deviation for each variable were also reported in the table.

To assess discriminant validity, average variance extracted (AVE) was analyzed, which measures the amount of variance captured by a construct in relation to the amount of variance due to measurement error (Fornell and Larcker, 1981). The formula for calculating AVE is as below:

\[
AVE = \frac{\sum_\lambda_i^2}{\sum_\lambda_i^2 + \sum[\text{Var}(\epsilon_i)]}
\]

where \( \lambda_i \) is the factor loading of each measurement item on its corresponding construct, and \( \epsilon_i \) is the error measurement. A widely used criterion to assess discriminant validity is Fornell-Larcker criterion (Fornell and Larcker, 1981), which suggested that based on the corrected correlations from the CFA model, the square root of a construct’s AVE should be larger than the coefficient of correlations between the specific construct and other constructs in the model—that is to say, a latent construct should explain better the variance of its own indicator rather than the variance of other latent constructs. Therefore, the square root of each construct’s AVE should have a greater value than the correlations with other latent constructs. If that is the case, discriminant validity is established on the construct level. In Table 2, evidence is provided for the construct validity of the scales.

**Measurement Invariance Tests**

Since the survey questions pertained to different referent groups (herding group vs people in the same village), multi-group confirmatory factor analysis (MGCFA) was conducted using Mplus following procedures recommended by Byrne (2013). These tests provide evidence that the observed scale indicators/items under study measured the same theoretical constructs (latent variables or factors) across the two groups of the sample. Without established measurement invariance, comparative analyses do not produce meaningful results, and results of differences between groups cannot be unambiguously interpreted (Milfont and Fischer, 2015).

Firstly, a baseline model (Model 1) was established from each group without constraints imposed across the groups for configural invariance (i.e., pattern invariance test). Next, Model 2 examining metric invariance was tested by

---

**Table 2** | Zero-Order Correlations, Means, and Standard Deviations of Measured Variables for Both Herding Group and Village Group Participants.

| Variables | Herding group participants | Village group participants |
|-----------|---------------------------|---------------------------|
|           | PDN | PIN | GID | GO | BI | PDN | PIN | GID | GO | BI |
| PDN       | 0.85|     |     |     |     | 0.83|     |     |     |     |
| PIN       | 0.02| 0.92|     |     |     | 0.07| 0.91|     |     |     |
| GID       | 0.67| 0.79| 0.63| 0.73| 0.75| 0.80| 0.79| 0.75| 0.78|     |
| GO        | 0.60| 0.02| 0.74| 0.22| 0.75| 0.74| 0.02| 0.74| 0.75| 0.76 |
| BI        | 2.09| 2.11| 2.11| 2.11| 2.11| 2.11| 2.11| 2.11| 2.11| 2.11 |
| M         | 3.21| 2.71| 3.38| 3.89| 2.02| 2.89| 2.56| 2.99| 3.81| 2.41 |
| SD        | 1.15| 1.08| 1.02| 0.64| 0.65| 0.84| 0.94| 0.96| 0.55| 0.84 |
| AVE       | 0.72| 0.85| 0.63| 0.53| 0.57| 0.70| 0.83| 0.63| 0.45| 0.61 |

PDN, Perceived Descriptive Norms; PIN, Perceived Injunctive Norms; GID, Group Identity; GO, Group Orientation; Figures in bold presented on the diagonal are the square roots of the average variance extracted (AVE).

*aCorrelation is significant at the 0.01 level (two-tailed).

*bCorrelation is significant at the 0.05 level (two-tailed).
TABLE 3 | Fit Indices for Measurement Invariance Tests between the Herding Group Members vs Village Group Members.

| Variable                  | Model                      | $\chi^2$ (df) | CFI ($\Delta$CFI) | $\Delta$ | $\Delta$2 ($\Delta$df) | Comparison       |
|---------------------------|----------------------------|---------------|-------------------|---------|-------------------------|------------------|
| Perceived Descriptive Norms | Model 1: Configural invariance | 0.00 (0)      | 1.00              |         |                         |                  |
|                           | Model 2: Metric invariance  | 3.84 (2)      | 1.00 (0.00)       | 0.15    | 3.84 (2)                | Model 1 vs Model 2|
|                           | Model 3: Scalar invariance  | 4.80 (4)      | 1.00 (0.00)       | 0.62    | 0.96 (2)                | Model 2 vs Model 3|
| Group Identity            | Model 1: Configural invariance | 1.00 (3)      | 1.00              |         |                         |                  |
|                           | Model 2: Metric invariance  | 5.86 (2)      | 0.99 (0.01)       | 0.06    | 5.86 (2)                | Model 1 vs Model 2|
|                           | Model 3: Scalar invariance  | 8.15 (4)      | 0.98 (0.01)       | 0.32    | 2.29 (2)                | Model 2 vs Model 3|
| Group Orientation         | Model 1: Configural invariance | 22.21 (4)     | 0.99              |         |                         |                  |
|                           | Model 2: Metric invariance  | 25.92 (7)     | 0.97 (0.02)       | 0.30    | 3.71 (3)                | Model 1 vs Model 2|
|                           | Model 3: Scalar invariance  | 29.09 (10)    | 0.97 (0.00)       | 0.37    | 3.17 (3)                | Model 2 vs Model 3|

CFI, comparative fit index.

constraining the factor loadings to be equal across the two groups (i.e., weak measurement invariance test). Model 3 tested scalar invariance by constraining both the factor loadings and indicator/item intercepts equal across the two groups (i.e., strong measurement invariance test; Byrne, 2013). Results showed no significant changes in Chi-squares across the three models, indicating a satisfactory measurement equivalence across the two groups. This enabled us to compare mean scores for the underlying factors across groups in the later analysis. The results were reported in Table 3.

Reliability Assessment

Following the establishment of scale dimensionality, parallelism, and invariance, reliability was assessed via calculation of Cronbach’s alpha for each scale using SPSS v.25, with both the split data file based on the referent group (i.e., herding group vs village group) and the combined dataset. Hunter and Gerbing (1982) suggested that when establishing new measures, validity and reliability should be treated separately. Hence, it was necessary to establish the dimensionality of the scales before examining scale reliability.

In addition to Cronbach’s alpha, composite reliability (sometimes called construct reliability) was assessed as an indicator of internal consistency in scale items (Netemeyer et al., 2003). By measuring the total amount of true score variance relative to the total scale score variance (Brunner and SUβ, 2005), it serves as an indicator of the shared variance among the observed variables used as an indicator of a latent construct (Fornell and Larcker, 1981). Thresholds for composite reliability are up for debate, but as a general guideline (Fornell and Larcker, 1981; Netemeyer et al., 2003), composite reliability of the constructs should be higher than 0.7; The formula (Netemeyer et al., 2003) is:

$$\frac{\left(\sum_{i=1}^{p} \lambda_i\right)^2}{\left(\sum_{i=1}^{p} \lambda_i^2 + \sum_{i=1}^{p} V(\delta)\right)}$$

where: $\lambda_i$ = completely standardized loading for the $i$th indicator, $V(\delta_i)$ = variance of the error term for the $i$th indicator, and $p$ = number of indicators.

Results (see Table 1) showed that coefficient alphas ranged from 0.60 to 0.93. Considering the uniqueness of the target culture group in this study and the fact that this was the very first study ever in which the measures were developed, the relatively lower-alpha scores for group orientation ($\alpha = 0.68$) and behavioral intentions ($\alpha = 0.60$) suggest that future use of these scales should correct estimates for unreliability due to error of measurement. The composite reliability estimates ranged from 0.77 to 0.91, providing additional evidence for scale reliability.

Ground Truthing Results

Step 7 in the MC3M is “ground truthing” of process, method, and findings throughout the entire course of the research with stakeholders, including cultural insiders. In the current study, this was accomplished in several key ways. First, by conducting cognitive interviewing and ongoing data and procedural quality checks during the course of the study, we accounted for perceptions of cultural insiders. Second, we regularly presented our procedures and progress to our community collaborators and enumerators to gain their input; changes to procedures were made when possible without compromising study rigor or validity. Third, the findings of the study were presented to people working in this region and on these topics prior to publication to discuss the findings and learn about their understanding of the study findings relative to their experience. Fourth, our project partners who work in this region and one of whom is a member of the population from which we sampled, were included in all publications and reviewed the content for consistency with their experience and understanding of the cultural context.

DISCUSSION

Noting the critical role of reliable and valid culturally derived measures for social norms constructs and the lack of models for developing measures in cultural context, the present study was designed to propose and apply a model to guide intercultural and cross-cultural communication researchers developing quantitative measures of study constructs. Specifically, this study contributed to the existing corpus of communication literature by offering the Model for Culturally Contextualized Communication Measurement (MC3M) to describe the process...
of developing measures for communication research involving unique populations. This model, derived from prior research in disciplines outside of communication and applied over several years in a program of research among ethnically Tibetan pastoralists, provides a clear path forward for researchers conducting studies of communication processes across or within cultures among marginalized or co-cultural groups. In addition to proposing and applying the MC³M, the results of this study provide preliminary evidence for measurement validity and reliability of measures of key social norms constructs. We first discuss the measurement development and findings using the MC³M process and then describe the utility and limitations of the MC³M.

Social Norms Measures
The development of the culturally contextualized measures of social norms constructs began with significant informal and formal information gathering processes and data collection. Existing social norms theories and measures (e.g., Cialdini et al., 1990; Lapinski and Rimal, 2005) and the culturally-contextualized conceptual definitions served as the basis for new item development and testing using a cross-sectional survey. The content evaluation was conducted by discussions among the multi-lingual, multi-cultural team members, translation and back-translation, and through cognitive interviews among participants from the study population. As a result, we modified questions, revised the response scale, and decided to use finger-counting as a way to describe the response scale to respondents. Continuous process and data quality monitoring during data collection contributed to the development of the measures.

Confirmatory factor analysis (CFA) provided initial evidence for the construct validity of the culturally derived social norm measures. Tests of internal consistency and parallelism indicated that the data were consistent with unidimensional factors measuring the two types of norms: descriptive norms and injunctive norms, as well as group identity, group orientation, and behavioral intentions. Notably, several items were removed from the scales for each of these constructs due to insufficient factor loadings suggesting the need for continued scrutiny of these items in future research. The items designed to measure outcome expectations failed to meet a priori standards, and as such, these items were removed from the final measurement analysis. Outcome expectations play a key role in enhancing the effects of social norms (Chung and Rimal, 2016), and future research should consider improved measures of this construct appropriate to cultural context. The failure of these items is difficult to explain. The content of the items was derived from in-depth interviews, and the adoption of procedures described by Ajzen et al. (1995) for belief elicitation was included; the item administration followed the same procedures as other scales. Nonetheless, it is clear that the items appear to be measuring unique concepts and do not form a unidimensional scale.

Most of the scales exhibited reliability coefficients within generally accepted ranges. However, the scale measuring behavioral intentions is relatively low. Perhaps this is due to the small number of items measuring this dimension since alpha is a function of the number of items on a scale. Because of the study procedures and the need to keep the questionnaire to a reasonable length to recruit and retain study participants without incentives, minimal items per dimension were administered. The behavioral intention scale could benefit from additional item refinement in future research studying behaviors in a cultural context. As an important limitation: although we focused a great deal on identifying, conceptualizing, and understanding the behaviors under study in the in-depth interviews (Lapinski et al., 2021), we did not focus our efforts on understanding our study community’s thinking about the concept of “intent.” This is something any legal scholar will remind us is complicated and perhaps culturally bound.

Because of the novelty of the study issue and information from our collaborators that most of our participants would not have the experience participating in research studies, a significant amount of time was spent reviewing and refining the item response scales. Ultimately, we decided to use digit counting and verbal descriptions of the responses. A “not sure” category was included in the scales, based on the cognitive interviewing process, and many participants used this option. The fact that many used this response option reinforces the importance of including it, but also makes the analysis and treatment of “not sure” responses complicated. It stands as a key limitation to our measures and will be explored carefully in future research. Reviewing the measurement literature for advice on how to handle these data, there was surprisingly little guidance. This represents an opening for future research on measurement and the development of response scales to be used when verbal administration of items is necessary, and populations may have little experience participating in research. This finding also highlights the utility of using cognitive interviewing to refine response scales and items.

Substantively, the “not sure” responses show that participants who were asked about village group members as the referent were more uncertain about what is considered normative behavior compared to those belonging to a herding group. These findings were consistent with the existing social norms and communication theories (e.g., Kincaid, 2004; Lapinski and Rimal, 2005; Mackie et al., 2015) on the critical role of physically or psychologically proximal groups in shaping, communicating, and maintaining normative information of certain behaviors.

Model for Culturally Contextualized Communication Measurement (MC³M)
The process described for developing, evaluating, and validating the culturally derived social norm measures presented in this study has valuable empirical and theoretical implications for researchers who intend to conduct studies of co-cultural groups or unique populations. The model delineating the specific steps in developing culturally derived communication measures,
starting from identifying and refining culturally derived conceptualizations, is a major contribution of this paper. Although we focus specifically on social norms research among the Tibetan population, we believe this model may have relevance for other communication research issues targeting other populations.

The MC3M has a number of key benefits and limitations. First, it provides a roadmap to researchers who wish to combine qualitative and quantitative methods to study communication processes in cultural contexts by specifying a set of best practices for developing measures. It is particularly applicable for populations or issues with little existing communication research, such as what we describe here. Second, it is based on existing research and practice and meant to function as a nascent and evolvable model as research on measurement development in cultural context progresses in the field of communication. There are certain additions and changes that could be incorporated into this model, and it is the hope of the researchers that it will have heuristic value, evolving as new knowledge is generated. Third, it is directly designed to be applied to intercultural, cross-cultural, and global communication research, filling a gap in the literature that has been dominated by other disciplines.

The model is not without limitations. Most importantly, we recognize that implementing the entire model requires significant time, resources, and relationships in a community. Further, the measures developed using the model cannot be simply taken and used in other cultural contexts but can serve as a basis for adaptation in intercultural communication research among similar populations and for similar issues. The relativism approach taken in the MC3M represents a departure from some of the existing cross-cultural/intercultural research, in which absolutism or universalism approaches are commonly adopted, and measures are used in communities without adaptation. With this said, we acknowledge that absolutism or universalism may still be appropriate in certain study contexts, such as when the research constructs are likely to be less sensitive to the influence of cultural or social factors.

Nonetheless, it is crucial to recognize the substantial role of culture in people’s communication, cognitions, and behaviors (Herdman et al., 1997; Berry et al., 2002). As such, we encourage researchers to develop quantitative measures derived within a specific cultural context following rigorous procedures. Measurement development and validation are critical for expanding social norms and other communication research accounting for cultural similarities and differences. Doing so can enhance both internal and external validity in the corpus of research to account for culturally-based concepts and processes (Mollen et al., 2010; Croucher et al., 2019).

The continued increasing global interactions highlight the need for cross-cultural researchers to be particularly careful and attentive to the issues of adapting existing constructs, theories, and measures developed in one culture for use in other cultures, and such issues are applicable to a variety of research disciplines. Acknowledging that nuances of the research process are different for each study, we hope that the proposed Model for Culturally Contextualized Communication Measurement, as well as the case we have described in this study, could serve to stimulate advancement in both conceptual and measurement refinement in intercultural and cross-cultural communication research.

### DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

### ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Michigan State University. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

### AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the many years of collaborative work and approved it for publication. The authors would like to acknowledge Ariane Leclerq, Ed Glazer, and the team of interviewers, coders, and surveyors/ enumerators for their assistance with this project.

### FUNDING

This project would not have been possible without a grant from the Sustainable Michigan Endowed Project at Michigan State University # 2011001. Partial support was provided by the USDA National Institute of Food and Agriculture, Hatch project numbers MICL02244, MICL02173, and MICL02362, and by National Science Foundation Award #SMA-1328503.

### SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcomm.2021.770513/full#supplementary-material
