Emotion between universalism and relativism:

Finding a standard for comparison in cross-cultural emotion research

Johnny R. J. Fontaine

Ghent University, Belgium

Seger M. Breugelmans

Tilburg University, the Netherlands

To be cites as: Fontaine, J., & Breugelmans, S. (2021). Emotion between Universalism and Relativism: Finding a Standard for Comparison in Cross-Cultural Emotion Research. In M. Bender & B. Adams (Eds.), *Methods and Assessment in Culture and Psychology* (Culture and Psychology, pp. 144-169). Cambridge: Cambridge University Press.

doi:10.1017/9781108675475.009

Emotion between universalism and relativism:

Finding a standard for comparison in cross-cultural emotion research

Emotion and culture share a long history. Perhaps one of the first empirical, cross-cultural studies was on emotion. Charles Darwin (1872/1998) published "On the expression of emotion in man and animals", analyzing within and between species continuity of emotion expression. As part of his studies he circulated queries to informants around the globe with respect to actual observations of emotion expressions. He received thirty-six reports, from which he concluded that "the same state of mind is expressed throughout the world with remarkable uniformity; and this fact is in itself interesting, as evidence for the close similarity in bodily structure and mental disposition of all the races of mankind." (p. 24). This conclusion is interesting because Darwin used his data to draw conclusions not only about the nature of emotions - in terms of specific expressions for specific mental states - but also about cross-cultural variation - in terms of universality. While this double focus of cross-cultural emotion research may have contributed to its popularity, it may also have contributed to a number of unresolved issues that hinder the development of a cumulative science of culture and emotion.

There have been, and still are, strong theoretical debates on the extent to which emotions are similar across cultures (i.e., present in similar form and function in populations around the globe), or rather that they are culture-specific (i.e., different in form and/or function across populations). These are generally known by the name of the universalism-relativism debate (Manstead & Fischer, 2002). Historically, a pendulum movement can be observed between both sides of the debate, with scholars sometimes being more in favor of the universalist position and sometimes more in favor of the relativistic position. For example, going into the 20th century, Darwin's universalist ideas were mostly replaced by views that emphasized the cultural variation and socialization of emotion (Birdwhistell, 1970; Klineberg,

1940). The case for universality was rekindled by perhaps the most well-known cross-cultural emotion research by Ekman and Friesen (1969, 1971) on facial expressions associated with the so-called basic emotions: happiness, anger, fear, disgust, sadness, and surprise (later complemented with contempt). On the basis of findings that most people on the globe recognize facial expressions in terms of the basic emotions they concluded that these are universal, hard-wired experiences. Later researchers again challenged this conclusion. For example, Russell (1994) questioned the criteria used for establishing universalism, claiming instead that there is substantial cross-cultural variation in recognition rates across cultural groups. Others, such as Lutz (1988) have challenged the notion that the six (or seven) basic emotions are universal, describing many cases of culture-specific emotions that derive their meaning from the specific cultural context.

The long history of cross-cultural emotion research does not mean that there is consensus on the basic findings. Universalism-relativism debates have proven to be quite resistant to resolution by data. In fact, different positions have been argued on the basis of the same data (e.g., Ekman, 1994 and Russell, 1994). Although reviews of the cross-cultural literature on emotions clearly point to evidence for both similarities and differences in emotions across cultures (e.g., Mesquita & Frijda, 1992; Mesquita, Frijda, & Scherer, 1997), it has proven too hard to "solve" the universalism-relativism debate and move beyond dichotomies, in part because there is little consensus on what data would constitute evidence for either position. In other words, the source of disagreement is often not the data as such, but rather the interpretation of what those data mean, either for emotion specificity or for cultural specificity (see Berry et al., 2011). This has led to a situation where debates on culture and emotion often lead to a polarization of opinions, without a prospect of conciliation or convergence (cf. Poortinga, 1997).

We argue that this state of affairs is undesirable. For a cumulative science of culture and emotion, different approaches need to be confronted and integrated. This means that we need to move from a situation of too much confirmatory orientation, with researchers of different positions working mostly within their own framework and premises, to a situation where specific expectations about cultural variation in emotions, both universalist and relativist, are formulated in relation to each other. In this paper, we forward four recommendations that are derived from the methodological and theoretical literature on emotion and on culture that, we hope, facilitate integration and exchange. The first two recommendations stem from the emotion literature and focus on making explicit what about emotions one is investigating and studying the meaning of the verbal emotion information that is used in the empirical studies. The last two recommendations apply theoretical and methodological reflections on cross-cultural differences to psychological constructs, some of which are also discussed in other methodological chapters in this book (e.g., chapter 14 and 15), to the study of emotions and invite to go beyond the relativism-universalism dichotomy. All four recommendations contribute to the development of a standard of comparison, a *tertium comparationis*, that should allow for a cumulative and constructive exchange between culture and emotion approaches.

1. Specify which facet of emotion you are investigating, preferably using a multicomponential approach

The first recommendation might seem so obvious that it would not merit a special mentioning; however, it might be one of the most important tasks for any emotion researcher and it is done much less frequently than might be expected. Like many latent psychological constructs, emotion is a fuzzy construct (Scherer, 2005). Most people intuitively know what feelings or emotions are, yet to scientifically define emotions in a manner that most scholars

would agree upon has proven to be a daunting task. Absence of consensus on what emotions are, is often a source of misunderstanding or even conflict between emotion studies; cross-cultural studies of emotion are no exception.

For cross-cultural research on emotions, such diversity in approaches means that studies can be hard to compare and that data of one study may not be generalized. For example, suppose that one researcher reports cross-cultural invariance in the extent to which specific facial expressions are characterized in terms of a given set of emotion terms (Ekman, 1994), while others report both cross-cultural invariance and cultural variations in subjectively experienced emotion (Scherer & Wallbot, 1994), and still another researcher reports substantial cultural variation in the terms of the emotion lexicon (Lutz, 1988; Jackson et al., 2019). How should we compare these findings with respect to the question to what extent emotions are culturally variable? The answer to that cross-cultural question in part depends on the question to what extent facial expressions, subjective experiences, and the emotion lexicon can be considered to be equivalent indicators of the same, latent emotion constructs.

The field of emotion has seen a continual movement on the question what an emotion is and how it should be measured. Whereas Wundt (1896) saw subjective feelings as the core of emotion, the theories by James (1884) and Lange (1885/1922) put psychophysiological changes at the center of emotion. Later approaches favored cognitions (more specifically appraisals; Schachter & Singer, 1962), action readiness (Arnold, 1960), facial expressions (Ekman & Friesen, 1971) or social symbols and roles (Averill, 1980). Each of these theories have their own preferred methodology to operationalize and measure emotions, focusing on the facet of emotion they deem to be central. For instance, theories that put subjective experiences central have focused on emotion lexicons and have identified up to four major dimensions in the experience of emotions (valence/hedonic tone, power/dominance,

arousal/activity, or novelty/unexpectedness, Fontaine & Scherer, 2013), while basic emotion theories that deem emotion expressions to be central have typically investigated how well people can recognize facial expressions of these basic emotions (e.g., joy, sadness, anger, fear, surprise, and disgust, Ekman & Friesen, 1971). More detailed analyses can be found in, for example, Frijda's (1986) standard work on emotion.

There is no consensus as of yet with regard to the best, or most optimal conceptualization of emotion, although the current literature appears to revolve around the ideas of (1) emotions as a combination of core affect (i.e., valence and arousal) and situated cognitions (Barrett, 2006) and that of (2) emotions as multi-componential phenomena that represent a synchronization of appraisals (the cognitive interpretation of the event), action tendencies (the motivational changes), bodily reactions (psychophysiological changes), expressions (through the face, the voice, or gestures that can be noted by others), and feelings (which represent the subjective awareness of the ongoing emotion process; Scherer, 2009). Both ideas see emotions as multi-componential phenomena but differ with respect to the status of the components; are these elements that get situationally connected with core affect (i.e., valence and arousal), or are these intrinsic elements of the emotions themselves?

Without committing to a particular emotion theory, the componential emotion approach can function as an overarching framework that allows to specify what about emotions one is investigating. The componential emotion approach allows to conceptualize the emotion domain as a grid of emotion processes, whether or not labelled by an emotion term, and emotion features which describe how these emotion processes are defined by or related to the different emotion components. One can focus on a single emotion component across emotions in the emotion domain, like studies that investigate core affect across emotions or studies about facial expressions of emotions, or one can focus on specific emotions and how they are characterized by one, more or all of these components. Depending

on the approach one takes, the challenges are different. When one focuses on a single component across emotions in the emotion domain, the question becomes how what is found for that one component generalizes to other aspects of the emotion process. When one focuses on a specific emotion, one has to demonstrate how all of the components make up or relate to that emotion and can be differentiated from related emotions in terms of these components.

Choosing the right measure then becomes a matter of what question about emotion one wants to address. For example, if one is interested in whether people have positive or negative feelings, then this suggests a dimensional approach testing for valence or evaluation (i.e., positive - negative). This dimension emerges from a wide variety of measures, most often from a set of emotion words (Yik, Russell, & Barrett, 1999). The attractive feature of this dimension for cross-cultural research is that it emerges even from culture-specific evaluations (Osgood, May, & Miron, 1975), so it is quite robust. The less attractive feature is that it tends to be a very general measure of emotion that harbors many possible specific effects (e.g., anger, fear, and sadness are all negative emotions but may have very different behavioral effects) and thus is not strongly predictive of most behaviors. The alternative is to go for a more specific emotion measure, such as a measure of shame, of disappointment, or of anger. There are various strategies to measure specific emotions. There are a few scales, such as the Test of Self-Conscious Affect (Tangney, 1990) that use vignettes, and others that use emotion components such as appraisals, bodily sensations, or action tendencies associated with the emotion. There are also non-verbal scales such as manekin approaches (Bradley & Lang, 1994) or analyses of facial expressions such as in the Facial Action Coding System (Ekman & Friesen, 1978).

The main challenge with specific-emotion measures is to show what specific emotions, or what specific characteristics about emotions, are measured in terms of convergent and discriminant validity. Emotions naturally show overlap in their features (i.e.,

specific emotion components; e.g., Frijda, Kuipers, & ter Schure, 1989; Roseman, 1991), so if a researcher wants to study the effects of a specific emotion, say guilt, she should show that the effects are due to that specific emotion and not to other emotions, say regret and shame, that share many features with guilt (e.g., Breugelmans et al., 2014; Fontaine et al., 2006). Many studies therefore also include features of these adjacent emotions to empirically verify whether effects are driven by the target emotion (discriminant validity). Similarly, if the valence of emotional experiences is measured, it would be good to also show that other dimensions that are repeatedly found in such studies, namely power/dominance, arousal/activity, or novelty/unexpectedness (Fontaine & Scherer, 2013), do not explain the data. This strategy is applicable to nearly any operationalization of emotions such as facial expressions, where one could either show that the pattern of action units for the target emotion is distinct from that of other emotions, or that the facial expression is categorized in terms of the target emotion and not in terms of another emotion.

Using a componential approach has been proven to be very fruitful in various cross-cultural studies. One early example is the large-scale study in 37 countries by Scherer and Wallbott (1994), who assessed patterns of emotion components for seven major emotions. At a general level, they found substantial evidence for cross-cultural universality of differential emotion patterning - the distinct profiles of emotions across the emotion components showed a large degree of cross-nation similarity. At the same time, they found notable cross-cultural differences in specific emotion elicitation, regulation, symbolic representation, and social sharing components. Componential approaches have also enabled the comparison of emotions for which no equivalent emotion terms could be found. For example, Breugelmans and Poortinga (2006) found that the Rarámuri, an indigenous group in Mexico, did distinguish in their reactions - in terms of emotion components - between the emotions of shame and guilt, even though they did not have distinct terms for these emotions. At the same time, they could

identify several emotion components that functioned differently with the Rarámuri (and with an Indonesian sample) compared to an international student sample, some of which could be interpreted in terms of meaningful differences (see Wong & Tsai, 2007). In a similar vein, Van de Ven, Zeelenberg, and Pieters (2009) demonstrated with a componential approach that experiences of two types of envy in Dutch (malicious envy and benign envy) could also be found in English and Spanish speaking samples even in absence of specific terms for these emotions in the respective languages.

To summarize, in order to create cross-cultural studies of emotion that are cumulative, it is imperative that researchers define how they conceive of and measure emotions in their studies. Unless researchers a priori have a specific measurement in mind, we propose that emotions are measured using a multicomponential approach. Besides integrating very different theoretical emotion approaches into a common conceptual framework, it can guide the researcher to representatively operationalize all key facets of the emotion(s) under study. It thus allows for a much more fine-grained assessment of the emotion constructs. With such a more representative operationalization of the emotion construct(s) a simultaneous assessment of universal and culture-specific aspects of emotion(s) is possible and allows to go beyond the simple dichotomy of relativism-universalism. Unless researchers are only interested in the dimensions underlying emotion, such as valence or activation, we strongly recommend against the use of emotion words as the sole measure of emotions.

2. Check whether emotion vocabulary is equivalent across languages and cultures

People use emotion words or phrases to express their feelings; so do researchers. We tend to use everyday vocabulary when we write or talk about studying affect, feelings or emotions. This poses a challenge to cross-cultural researchers because the equivalence of the emotion vocabulary between one language and another cannot be assumed. This has

consequences, not only for the stimulus and response materials that are used, but also for the theoretical position that is taken with regard to emotions in general.

The importance of the emotion vocabulary for the psychological process(es) is still debated, echoing the old Sapir-Whorf discussions on linguistic relativity. There are researchers who see language mostly as a vehicle for communication, arguing that feelings and emotions can exist independently of the available emotion vocabulary (Frijda, et al., 1995). There are others who see the relationship as being more intimate, arguing that the available emotion vocabulary shapes the experience of core affect (i.e., valence and arousal) into the discrete emotions we generally talk about (Barrett, 2006). As cross-cultural emotion researchers it is good to be aware that other researchers may hold quite varied opinions with regard to linguistic relativity and that issues that may appear self-evident from one's own perspective (e.g., "emotion words are just labels that we use to express our feelings; what we feel is independent of these words" or rather "different meanings of emotion words across cultures are an expression of a deeper, cultural construction of the emotions themselves") are likely not shared among peers. This is all the more reason to pay explicit attention to the emotion and emotion-related words that are used. The topic of linguistic relativity is vast and still debated, so we will not take a position in this chapter. Rather, we want to address the challenges that arise from a more practical point of view when studying emotions by means of words across linguistic groups.

Much emotion research uses verbal material either as stimuli or response items, for example in mood questionnaires where emotion and feeling terms are rated in terms of intensity (e.g., "How strongly do you feel anger?"). When this verbal material has a different meaning in the different cultural groups, incomparability of data is built-in into the research and any comparison between the different groups in terms of the emotion construct is invalid. After all, both groups in effect responded to different stimuli or response items. For example,

Osgood, May and Miron (1975) found that the term “aggressive” was unique in having a slightly positive connotation in American English. Thus, any literal translation of “aggressive” in closely languages such as French (*agressif*) or German (*agressiv*) would actually measure different constructs. Another example is the research on facial expressions (Ekman & Friesen, 1971). The goal of this research was to see whether specific facial expressions are recognized in terms of specific emotions. An often-used method involves participants seeing a picture of a facial expression and then identifying which emotion this expression conveys by choosing from a list of six emotion words, representing the basic emotions (e.g., happiness, sadness, anger, disgust, fear, surprise). Thus, the choice of an emotion word is taken to be an indicator of the emotion. Of course, if emotion words differ in meaning across cultures, this assumption becomes questionable. Indeed, as Russell (1994) discussed, absolute recognition rates did show a lot of cultural variation, with recognition in Asian and African societies being markedly lower than in Western-European and American samples, sometimes as low as 46% for fear and 50% for anger in Malay samples. One reason for these differences could be that the emotion words that are used simply differ in meaning. Russell saw these findings as a challenge to the idea of universal, basic emotions, while Ekman (1994) saw linguistic differences as a source of error that only strengthened the finding that in most cultures the "right" word was matched with the "right" emotion at a rate higher than would be expected on the basis of chance. Irrespective of which interpretation is more plausible, it does illustrate the importance of checking the meaning of emotion words used in a study.

Knowing that substantial differences in both the availability of terms and the meaning of terms may arise, it is wise to address the meaning of emotion words explicitly in research. As was mentioned under the previous recommendation, using multiple components as indicators of an emotion allows for a much more fine-grained analysis of equivalence of

emotions. It is important to identify potential meaning shifts in the emotional vocabulary on beforehand. Careful translation procedures, which nowadays often include a combination of translation-back-translation and committee approaches, can identify the most glaring problems (Van de Vijver & Leung, 1997). However, even careful translation processes do not guarantee that the translated material has the same meaning across language groups.

A special word should be said here about methods that do not use emotion words as part of the stimulus or response materials. Although they are used relatively less frequent than verbal emotion measures, they still cover a substantial part of the emotion literature. Apart from the facial expressions by Ekman measured by the Facial Action Coding System (FACS; Ekman & Friesen, 1978) or by software such as FaceReader (Loijens & Krips, 2018), examples are measures using pictures (e.g., the IAPS, Lang, Bradley, & Cuthberth, 1999; or the NAPS, Riegel et al., 2017), body postures (e.g., Kret, Stekelenburg, Roelofs, & de Gelder, 2013; Van Osch, Zeelenberg & Breugelmans, 2016), or emotion manekins (e.g., Bradley & Lang, 1994). In addition, there are a host of psychophysiological measures such as EEG and EMG, galvanic skin response or pupil dilation that are used to measure emotion, albeit few of such studies address cross-cultural issues (e.g., Levenson, Ekman, Heider, & Friesen, 1992).

While it is true that the use of such non-verbal measures sidesteps many of the issues we note with emotion vocabulary, they only do so partially. For one, many of these measures have been validated by comparing the non-verbal stimuli to verbal ratings. The best-known example are the facial expressions of the basic emotions which were classified by a forced-choice method in which participants chose which of six verbal labels best describes the emotion expressed. This re-introduces verbal biases when doing a study in different linguistic communities (see Russell, 1994 for an elaborate discussion). The same validation issues count for most other non-verbal measures as well. In addition, there is the issue of interpretation. Responses are seldom interpreted as they are (e.g., galvanic skin response), but rather are

described by researchers in verbal emotion categories, often derived from the English language, such as anger, fear or sadness. This interpretation in terms of emotion labels again introduces the possibility of bias in the meaning of emotion vocabulary. To conclude, while non-verbal measures surely suffer less from meaning differences in emotion words, they are by no means immune to them. Thus, it is wise to directly assess the meaning equivalence of emotion words, also when they are only used to interpret non-verbal assessments. Here, we would like to draw attention to a new psycholinguistic assessment method that allows for the assessment of the meaning equivalence emotion words. This method has been developed based on the componential emotion approach and is called the GRID method (Fontaine et al., 2007; 2013). It consists of a simple grid of emotion terms and emotion features that represent each of five emotion components. In its basic version, the GRID instrument consists of 24 emotion terms and 142 emotion features. The 24 emotion terms were selected on the basis of which emotions received most scientific attention, which emotion terms were reported spontaneously in a survey of the Swiss household study, and which emotion terms were frequently reported in cross-cultural free-listing studies. The 142 features were selected to present five major components that make up or are strongly related to the emotion processes: appraisals, action tendencies, bodily reactions, expressions, and feelings. Importantly, this componential emotion approach allowed to use very different emotion models and theoretical frameworks to operationalize the features that characterize emotion processes. For instance, appraisals were operationalized on the basis of Scherer's sequential check model (1999), action tendencies on the basis of Frijda's action tendency model (e.g., Frijda et al., 1989), and feelings on the basis of Osgood's evaluation-potency-activation model (e.g., Osgood et al., 1975), which each theory having a very different approach to what emotions are. The information generated by the GRID is thus relevant for research based on very different emotion theories.

In the standard GRID, participants evaluate the meaning of a subset of four emotion terms with respect to the 142 emotion features in the cultural and language group to which they belong. Applying this instrument in 34 samples in 27 countries and 23 languages across the world with about 5000 participants, it was observed that all emotion components were involved in the meaning of emotion terms in all languages studied. Across the 24 emotion terms and the language groups a very stable four-dimensional structure emerged consisting of valence (negative versus positive; e.g., "feeling good" and "frowning", "joy" and "sadness"), power (weak versus strong; e.g., "firm voice" and "weak limbs", "sadness" and "anger"), arousal (high arousal versus low arousal; e.g., "decreasing heartbeat" and "increasing breathing", "disappointment" and "stress") and novelty (low novelty versus high novelty; e.g., "experienced the emotional state for a long time" and "jaw drop", "guilt" and "surprise").

Importantly, these overall similarities in dimensional structure could be used as a *tertium comparationis*, the quality that the things that are compared (i.e., emotions across cultures) have in common, to identify meaning differences between cultural and linguistic cultural groups, both with respect to the meaning of emotion terms and emotion features. For example, in higher power-distance groups “pride” was evaluated as less positive compared to low-power distance groups. As another example, the item "feeling alert" varied hugely in meaning with respect to the valence dimension, indicating that the alertness concept may contains two aspects: "being alerted" which represents a negative emotional experience and "being awake and attentive" which is characteristic for a positive emotional experience. Although translations were carefully executed, this shift went unnoticed during the translation process and was only revealed with the GRID instrument.

The findings with the GRID instrument are interesting in their own right. They both reveal stable properties in the meaning structure of the emotion domain across languages and cultural groups (in terms of the main structuring dimensions of the emotion domain) fitting a

universalistic approach, as well culture-specific variations in emotion meaning fitting a relativistic approach. The GRID method thus allows to go beyond the universalism-relativism dichotomy, and even more, reveals that both perspectives can be interdependent. It is just through identifying universal properties about the meaning structure, that culture-specificity clearly emerged.

The GRID instrument is also interesting from an applied perspective. The instrument and the extensive database already collected with this instrument can be used to empirically identify shifts in meaning, both with respect to the meaning of emotion terms as with respect to the meaning of emotion features that define the emotion process. By applying this instrument, it can be avoided that incomparability is built-in in the instruments that are used to assess cross-cultural differences in emotion. Moreover, empirically identifying patterns in the meaning shifts of emotion terms and in the meaning of features can generate focused hypotheses about the impact of culture on the emotion processes, as is for instance the case with shifts in the meaning of pride.

To summarize, the second recommendation is to ensure that the emotion terms that one is using, have an equivalent meaning in the various cultural groups. This means not just relying on translation equivalence, but rather empirically investigating the meaning of emotion terms and emotion features. We think the GRID instrument provide an accessible method to do so. Even if researchers prefer other methods, such as semantic differential analysis (Osgood, May, and Miron, 1975), Multidimensional Scaling of similarity ratings (Russell, 1980), or network analyses of colexification (Lange et al. 2019), checking for the equivalence of emotion terms is paramount.

3. Specify the nature of the cross-cultural comparison

Strong relativist approaches focus on qualitative differences in the processes and traits that govern psychological functioning across cultural groups. This means that, in order to understand how these culture-specific processes and traits function, one has to study the internal organization of each cultural system. An often-used example is that of the emotion of *liget* among the Ilongots studied by Rosaldo (1980). *Liget* is the emotion that the Ilongots experienced when participating in the cultural practice of headhunting. It is described as a cultural-specific emotion, explained by a specific aspect of Ilongot culture. Thus, from a strong relativist perspective, this emotion can only be studied within the cultural context they emerge. From a strong universalist approach, however, psychological traits and processes are biologically rooted and therefore universal. Cross-cultural (epi)genetic differences account for quantitative differences between cultural groups on these psychological traits and processes. The notion of affect as hardwired, biological mechanisms in Tomkins' (1962, 1963) affect theory (which formed one of the sources of inspiration of Ekman and Friesen's studies on the universality of the basic emotions) could be an example. Based on this approach, (aspects of) emotions can be directly compared between cultural groups and the possible impact of culture is very limited.

One of the reasons for the persistence of the universalism-relativism debate is that questions about cross-cultural variation are often posed in the above sketched dichotomous manner. An emotion is either the same across cultures (= universalism) or is it not (= relativism). This means that any study is either expected to be generalizable to all aspects of emotions and cultures, or not to be generalizable at all. That is not realistic. Emotions can be comparable or incomparable in different ways that do not fit this simple universalism-relativism dichotomy. Most researchers hold ideas on a continuum between universalist and relativist positions, however, without specifying what is comparable or different. We present a simple model, based on a series of three successive questions, to specify the nature of the

expected and inferred cross-cultural differences and/or similarities that goes beyond the simple dichotomy. It is important to clarify one's position here, because it entails specific assumptions, interests, and questions that have implications for research design and data analysis.

The three questions may help researchers in identifying their theoretical interests and positions as well as inform them about what they expect to be similar and what they expect to differ across cultural groups. The questions stem from previous, fine-grained conceptual frameworks and methodological models on relativism-universalism (Berry et al., 2011, Fontaine, 2011), on bias and equivalence (Fontaine, 2005; Van de Vijver & Tanzer, 1997), and on multilevel isomorphism (Fontaine, 2008; Van de Vijver, van Hemert, & Poortinga, 2008).

The first question is whether *a psychological domain is expected to be organized according to the same underlying traits and processes in each of the cultural groups*. To borrow an analogy from the personality domain: do the Big Five personality traits of Extraversion, Neuroticism, Conscientiousness, Agreeableness, and Openness describe personality differences within all cultural groups of the world? In a similar vein, we could ask whether the same dimensions of affect (e.g., evaluation, activation, potency) or whether the same organization of emotion components account for emotional behavior in each of the cultural groups across the world? If the answer to this question is no, we are in a scenario of *full cultural specificity*. This means that the psychological domain of interest, in our case emotions or a subdomain of emotions, can only be conceptualized and empirically studied within each cultural group separately. Psychological insights do not cross borders in this scenario, as in the previously reported example of *liget* which only makes sense within the Ilongot's culture. We need to resort to within-culture descriptions and, at most, interpretations

of different within-culture descriptions (cf. the Human Relation Area Files (http://hraf.yale.edu/).

If the answer to the first question is yes, however, a second, subsequent question can be asked: *Do these universal traits and processes emerge in the same observable behavioral repertoire in each of the cultural groups?* It can be that a psychological process is shared across cultural groups, but that it emerges in very different behavior. For instance, all cultural groups display greeting behavior, but the exact behavior is different (such as hand shake, knotting, bowing, mutually touching noses). With respect to emotions, it has for instance been suggested that shame is expressed very differently in certain regions of India (Menon & Shweder, 1994) compared to more Western countries; as seen in the culturally well-known expression by the Goddess Kali, shame can be expressed by a tongue protrusion (biting one's tongue), which is not part at all of the expression of shame in Western groups. Thus, while the emotion is the same, the way it emerges in the actual behavioral repertoire can be very different. If the answer to the second question is no, we have a scenario of *functional equivalence*. In this case, a common theoretical framework can be developed to account for a psychological domain, but the actual behavior that is studied and used for assessment in each group will differ and will thus not allow for direct quantitative comparisons between cultural groups. Qualitative comparisons, like the one described above on shame, are very well possible.

The difference between the full cultural specificity scenario and the functional equivalence scenario is not situated in what is precisely observed (both agree that the observed behavior is different), but in how the behavior is accounted for (whether or not by universal underlying traits and processes). There are basically two ways to empirically demonstrate whether or not different phenomena can be accounted for by the same underlying theoretical variables: studying construct representation and nomological networks, which are

also two main sources of validity for psychological assessment instruments (Messick, 1989; APA Standards for Educational and Psychological Testing, 2014). With construct representation, the direct processes that lead to the target behavior are studied and compared. In the emotion domain, the componential emotion framework lends itself very well to the application of a construct representation approach, as it specifies which emotion components make up or are closely related to the emotion process one is investigating. One can identify to what extent the same relationships between emotion components emerge across cultural groups. For instance, one could investigate whether the same appraisals of being under public scrutiny (key appraisal for shame) and having a tendency to disappear from the social interaction (key action tendency for shame) are expressed by a tongue protrusion in India. A second way to demonstrate or reject the claim of a similar theoretical construct that accounts for different observable behavior is to study the nomological network. When it can be demonstrated that the relationships with antecedent and consequent constructs are the same across cultural groups, the same theoretical variable can be assumed to account for the culturally different observed behavior.

If the answer to the second question is yes, a third question follows: *Are quantitative differences between groups in the behavioral repertoire organized in the same way as differences between individuals within groups?* This third question has been raised by multilevel approaches in the social sciences (e.g., van de Vijver & Poortinga, 2002). It has been noted that there is independence between levels in hierarchically organized structures (such as pupils in schools). The factors that predict individual pupils' achievement can be very different from the factors that predict how well a school achieves as a whole. Cross-cultural and cultural psychology are par excellence characterized by a hierarchical structure: people are a member of a cultural group and do not easily change membership. This means that

cultural differences in psychological characteristics need not have the same structure as individual differences and thus need not be necessarily interpreted in the same way.

If the answer to the third question is no, we have a scenario of *behavioral repertoire equivalence*. In this scenario, the same conceptual framework can be developed and the same behavior can be studied and used for assessment for within-cultural purposes, but individuals from different cultural groups, or the cultural groups themselves, cannot be directly compared. Only the within-cultural differences between individuals can be accounted for by the same psychological constructs. What happens at a cultural level has to be accounted for by different constructs. A very clear example from the value domain are the values "being humble" and "social power". Within cultures, these values tend to be negatively correlated and are considered indicators of two different value types, namely the tradition and the power value type (Schwartz, 1992). However, at a cultural level, both values are positively correlated and can be both seen as indicators of the amount of hierarchy valued in a culture (Schwartz, 2006). Similarly, in the affective domain it has been found that at an individual level, divorce is a typical antecedent of negative emotional experiences (e.g., Sbarra, 2015). At a cultural level, however, the proportion of divorces is not at all correlated with the average reported happiness in countries (http://www.worldvaluessurvey.org/WVSOnline.jsp). Thus, the theoretical constructs to understand the emotional impact of divorces at individual and cultural level are different. A possible explanation is that at the individual level a failure to live up to a life commitment is an antecedent to a whole gamut of negative emotional experiences, while at the cultural level the proportion of divorces is also indicative for the freedom a society allows its members to choose for and change one's life commitments and which has at least partially a different meaning.

Thus, compared to the construct equivalence scenario, the repertoire equivalence scenario assumes that the same psychological constructs emerge by and large in the same way

in the behavioral repertoire across cultural groups. To demonstrate this type of equivalence it is important to both demonstrate that the domain of generalization of the construct in each of the cultural groups is highly comparable as well that the comparable behavior is characterized by the same internal structure in each of the cultural groups. To demonstrate the former, qualitative research approaches are indicated, like participant observation, interviewing, focus groups, working with local experts, and surveys with open qualitative questions (e.g., Fontaine, 2005; van de Vijver & Leung, 1997). A typical example from the emotion domain is the identification of culture- and language-specific emotion vocabularies (Fontaine et al., 2002). By using free listing tasks, were participants are free to write down as many emotion terms they can think of in their own language, is a good way to identify the emotion terms that are salient in a cultural/linguistic group. However, qualitative research is insufficient to show that the same psychological constructs effectively emerge in the same (or at least largely overlapping) behavioral repertoires. It must also be demonstrated that the same or similar behavior is characterized by the same internal structure. Data analytic methods that are typically used to compare the internal structures are exploratory or confirmatory factor analysis, multidimensional scaling, and recently suggested psychological network analyses (Lange et al., in press). Only when the psychological organization of the domain is the same, can the same observed behavior be used within each of the cultural groups. For instance, after a free listing tasks in The Netherlands and Indonesia, Fontaine et al. identified 50 translation equivalent emotion terms. Only for a subset of 42 terms they could demonstrate that they shared the same position in a three-dimensional valence, power, and arousal space. The other pairs of eight terms were well represented in the three-dimensional representation, but did not share the same location for the two groups, and thus shifted in meaning between the two groups.

If the answer to the third question is yes, we have the scenario of *isomorphism*. It has for instance been demonstrated that the Big Five personality structure holds both within cultural groups and between cultural groups. Only when there is isomorphism it is possible to study individual and cultural differences with the same theoretical framework and to directly compare individuals from different cultural groups with one another. This would be the case if all characteristics that typically define shame at the individual level (like appraising a situation a thread to one's reputation, getting warm, wanting to disappear, …) would all be more salient in shame cultures and that all characteristics of guilt at an individual level (like appraising oneself as responsible, being concerned with the negative consequences of one's behavior, intra-punitive behavior and reparative tendencies) are all more salient in guilt cultures. With respect to the dimensional model, Kuppens and colleagues (2006) have for instance demonstrated that both within and across cultural groups a positive and a negative affect factor represented emotional experiences well.

If (ideally) a random subset of cultural groups is studied, isomorphism can be investigated by means of multilevel analyses, such as multilevel factor analysis and multilevel regression analysis (e.g., Fontaine & Fischer, 2011). When only two or a few cultural groups are studied, isomorphism can be studied by investigating item bias (e.g., van de Vijver & Leung, 2011). Both non-uniform and uniform item bias can be studied with IRT models or with confirmatory factor analysis. Only when at least a subset of non-biased items can be identified, can direct comparisons between cultural groups be justified.

These three stepwise questions create a four-fold framework that does not just go beyond the dichotomy culture-specific/relativist versus culture-similar/universalist, it also specifies what aspect of the psychological construct is specific and what aspect is similar. Thus, depending on the scenario, cultural differences (1) encompass the whole psychological construct of investigation, in our case one or more emotions, (2) are situated in how

qualitatively universal emotions manifest themselves into a culture-specific repertoire (i.e., eliciting situations, relations with other constructs, or emerging behavior), (3) affect how (characteristics of) emotions are structured differently at the cultural compared to the individual level, or (4) consist of quantitative differences in emotions between cultural groups (implying that the individual and cultural level are organized in the same way). Since at each step in the four-fold model additional restrictions are imposed on what is comparable between cultural groups we can think of these four scenarios as four hierarchically ordered levels of culture-specificity/comparability.

To summarize, the third recommendation is to specify the level at which emotions are taken to be culture-specific/similar across cultures in your study. Doing this not only helps other researchers understanding the theoretical contribution of your research, it also makes clear where you expect (or assume) cross-cultural equivalence in your data and where not, and to empirically justify this claim. This, in turn, makes findings from different studies more easy to integrate and makes interpretations of data more easy to evaluate. Even if you would prefer a different framework, we strongly recommend that you stay clear of simplistic universalism-relativism dichotomies (it's all the same or it's all different) if you wish cross-cultural studies on emotion to be cumulative.

4. Account for both similarities and differences

Nearly all cross-cultural studies find evidence for both cross-cultural similarities and differences in their data. Similarly, all reviews of cross-cultural studies of emotion find evidence for both similarities and differences in emotion or components of emotion (e.g., Berry et al., 2011; Mesquita & Frijda, 1986; Mesquita, Frijda, & Scherer, 1997). Yet, something strange appears to happen when researchers interpret their data. In spite of finding both, they usually interpret only differences or only similarities (cf. Brouwers et al., 2004). It

is this one-sided interpretation, more than the actual data, that contributes to the universalism-relativism dichotomy. For a cumulative science of emotion across cultures, it is necessary that researchers account for all of their data, similarities and differences alike. In the previous recommendation, we described how different theoretical perspectives implied a different balance or source of similarities and difference. In this fourth recommendation, we address how to account for these similarities and differences.

There are at least two places where similarities and differences could be explicitly addressed. The first is in the expectations or hypotheses that are formulated. As we discussed under recommendation 3, any comparison between two or more cultural groups requires that there is something that is comparable (=similar) between groups. For example, if we expect people from honor cultures to react with more intense anger to insults compared to people from non-honor cultures (Cohen & Nisbett, 1997), then we assume that anger (and insults for that matter) have the same meaning in both cultures. At the same time, nearly any comprehensive study of culture and emotion is bound to find differences. Even studies that reported evidence for universality in emotions found a substantial number of cross-cultural differences (e.g., Breugelmans et al., 2005). This means that, if one's hypotheses are about similarities, it would be wise to see whether expectations about specific differences - even if only in specific emotion components or antecedents - could be made on the basis of previous literature.

The second place where similarities and differences should be addressed is in the interpretation of the results. Researchers sometimes draw conclusions that are more one-sided than their data warrant, emphasizing either the similarities or the differences, some even generalizing findings to claims of universalism or relativism of emotions as a whole. The universalism-relativism dichotomy is reinforced when research is cited for one-sided interpretations, rather than for the data, which usually are more nuanced. This does not mean

that researchers cannot have a preferred position, being interested most in either similarities or differences; it just means that both should be acknowledged. This recommendation extends to editors and reviewers. Though nuanced stories may appear less appealing or "sexy" than papers that drive a single point home, it is their responsibility to stimulate balanced interpretations of the data. As an illustration, Breugelmans et al. (2014) set-up a study to test for structural equivalence of emotion components associated with regret, guilt, and disappointment. They expected and found that nearly all components clustered into the three emotions in the same way in four national samples. However, they also found unexpected cross-cultural differences in the frequency and intensity of emotions that were very consistent with cultural theories about the importance of engaged and disengaged concerns (e.g., Kitayama, Markus, & Kurokawa, 2000). In the end, this meant that half of the results section consisted of analyses that were not mentioned in the introduction; this set-up was encouraged by several of the reviewers and the editor.

When explaining similarities and differences there appears to be a preference for different distal factors. Relativism and universalism are typically taken to coincide with cultural versus genetic explanations. However, this does not need to be the case. Both genetic and cultural factors could account for any of the four positions described in section 3, and most likely a combination of both is needed to account for the observed cultural similarities and differences.

Although genetic factors are typically proposed for similarities in psychological functioning and traits, qualitative cross-cultural differences could possibly be accounted for by genetic factors. Kitayama et al. (2014) have suggested that differences between cultural groups in self-concept have a genetic basis, with some genes being more prevalent in interdependent cultural groups and other genes in independent cultural groups. This would mean that cultural differences in engaged versus disengaged emotional functioning would

have a genetic explanation. Genes could also account for the existence of a psychological process or trait, without accounting for how it emerges in the behavioral repertoire. For instance, it could be argued that greeting behavior is a genetically predisposed appeasement behavior (cf. Darwin, 1872/1998), without the precise form of the appeasement behavior being genetically determined. In similar vein, it could be the case that shame is a genetically evolved mechanism to indicate awareness of social rule violation, without the precise shame expression being genetically determined (cf. Fessler, 2007). This means that demonstrating cross-cultural similarities in the expression of emotions, cannot be taken as evidence in favor of a genetic explanation (cf. Ekman, 1994). Effects where genetics are indirectly related to emotion may also be found. For instance, blushing will be more or less salient for the repertoire of shame depending on the genetically determined skin tone. The darker the skin, the less blushing can be noticed by others and the less it will play a role in the interaction with others (even though the actual increase in blood-flow is independent of skin tone). It also means that cross-cultural differences do not make a sufficient argument against a genetic account of emotion processes.

Cultural factors are typically used to account for cross-cultural differences; for example, if the psychological process is culturally created, as in the example of *liget*. Cultural factors could also account for the example about greeting. Through vicarious learning, people learn how to greet in a specific culture. As with genes, cultural factors are not necessarily tied to differences. For example, universal organism-environment interactions, cultural diffusion, blending or globalization might account for similarities in emotion and emotion expression.

Ideally, studies explaining similarities and differences in terms of genes or cultural factors would provide additional evidence that supports such explanations. Definitions of culture are so broad and multifaceted (Soudijn, Hutschemaekers, & Van de Vijver, 1990) that some have argued that it is not useful as a scientific construct for psychology at all (Poortinga,

2015). When trying to explain differences in terms of culture there are at least two promising strategies. One is the decomposition of cross-cultural effects into smaller, more specific effects (i.e., "peeling the onion", Poortinga, Van de Vijver, Joe, & Van de Koppel, 1987). For example, if people from a herding society (honor culture) are expected to react with stronger anger (Cohen & Nisbett, 1997). Another is the identification of a (psychological) process that can explain the observed differences (Matsumoto & Yoo, 2006). This is especially viable if the structure of emotions is the same between countries as within (question 3 from Recommendation 3). For example, if it is found that people from certain cultures react with more anger and shame to insults because of their honor values, then these values could be measured. If it is found that, within cultures, stronger honor values predict more anger and shame in reaction to insults and that, between cultures, the purported "honor" culture scores higher on honor values and on anger and shame in reaction to insults than the "non-honor" culture, this is a powerful support for a cultural explanation. For genes, the story is similar, though perhaps more complex from a practical perspective. Given that most complex traits that psychology studies, are caused not by a single gene, but rather by multiple genes, measuring the relevant alleles and showing that they are related to the target differences may be a daunting task, although the rapid development of measurement techniques may change that in the near future.

To summarize, the fourth recommendation is to explicitly address both similarities and differences in the introduction and in the discussion of the results. This avoids simplistic universalism-relativism dichotomies and allows researchers to specify what they think is interesting in the data, without ignoring part of those data altogether. Explanations of data in terms of genes or culture may change over time, as they have done in the past, but data should not change depending on who interprets them. A cumulative science of emotions across

cultures should let data prevail and these data in nearly all cases include both similarities and differences. Thus, they should both be explained.

Conclusions

In the current chapter, we have brought together insights from the emotion literature, from theoretical reflections on the nature of and accounts for cross-cultural differences and similarities, and from the methodological bias, equivalence and multilevel literature in cross-cultural psychology to propose four recommendations to bridge the gap between relativist and universalist approaches. To briefly summarize, we recommend that researchers:

1. specify the emotions or facets of emotions they study, preferably using a multi-componential approach to measurement;

2. check whether emotion vocabulary they are using is equivalent across languages and cultures, either by using existing data bases or by including a measurement of meaning in their design;

3. specify the level at which they are comparing emotions across cultures, ranging from description of culture-specific constructs to comparing mean scores between cultural groups and apply the adequate methods to demonstrate the level of comparability claimed;

4. account for both similarities and differences in their data, both in formulating their hypotheses and in interpreting their data.

If these recommendations seem obvious to the reader, we are very content because we share their opinion. It is our conviction that the field is ready to move beyond dichotomized debates and to move toward a cumulative science of culture and emotion, where by creating a conceptual and methodological framework data can be used to inform true exchange between theoretical positions. The cross-cultural study of emotion has a long and fascinating history. We believe it has at least an equally fascinating future.

References

Arnold, M. B. (1960). *Emotion and personality. Vol. 1. Psychological aspects*. New York: Columbia University Press.

American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational and Psychological Testing (U.S.). (2014). *Standards for educational and psychological testing.* Washington, DC: AERA.

Averill, J. (1980). A constructivist view of emotion. In R. Plutchik & H. Kellerman (Eds.), *Emotion: Theory, research, and experience* (Vol. 1, pp. 305-340). New York: Academic Press.

Barrett, L. F. (2006). Are emotions natural kinds? *Perspectives on Psychological Science*, *1*, 28-58.

Berry, J. W., Poortinga, Y. H., Breugelmans, S. M., Chasiotis, A., & Sam, D. L. (2001). *Cross-cultural psychology: Research and applications* (3rd Ed.). Cambridge: Cambridge University Press.

Birdwhistell, R. L. (1970). *Kinesics and context.* Philadelphia: University of Philadelphia Press.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25*, 49-59.

Breugelmans, S. M., Pootinga, Y. H., Ambadar, Z., Setiadi, B., Vaca, J. B., & Widiyanto, P. (2005). Body sensations associated with emotions in Rarámuri Indians, rural Javanese, and three student samples. *Emotion, 5*, 166-174.

Breugelmans, S. M., & Poortinga, Y. H. (2006). Emotion without a word: Shame and guilt among Rarámuri Indians and Rural Javanese. *Journal of Personality and Social Psychology*, *91*, 1111-1122.

Breugelmans, S. M., Zeelenberg, M., Gilovich, T., Huang, W. H., & Shani, Y. (2014). Generality and Cultural Variation in the Experience of Regret. *Emotion, 14*, 1037-1048.

Brouwers, S. A., Van Hemert, D. A., Breugelmans, S. M., & Van de Vijver, F. J. R., (2004). A historical analysis of empirical studies published in the Journal of Cross-Cultural Psychology - 1970-2004. *Journal of Cross-Cultural Psychology, 35*, 251-262.

Cohen, D., & Nisbett, R. E. (1997). Field experiments examining the culture of honor: The role of institutions perpetuating norms about violence. *Personality and Social Psychology Bulletin, 23*, 1188-1199.

Darwin, C. (1998). *The expression of the emotions in man and animals*. (P. Ekman, Ed.) (3rd ed.). London: HarperCollins. (Original work published 1872).

Ekman, P. (1994). Strong evidence for universals in facial expression: A reply to Russell's mistaken critique. *Psychological Bulletin, 115*, 268-287.

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*, *1*, 49-98.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in face and emotion. *Journal of Personality and Social Psychology, 17*, 124-129.

Ekman, P., & Friesen, W. V. (1978). *The Facial Action Coding System: A technique for the measurement of facial movement*. Palo Alto, CA: Consulting Psychologists Press.

Fessler, D. M. T. (2007) From appeasement to conformity: Evolutionary and cultural perspectives on shame, competition, and cooperation. In J. L. Tracy, R. W. Robins, &

J. P. Tangney (eds.), *The Self-Conscious Emotions: Theory and Research* (pp. 174-193). Guilford Press.

Fischer, R., & Fontaine, J. R. J. (2011). Methods for investigating structural equivalence. In D. Matsumoto & F. van de Vijver (Eds.), *Cross-cultural research methods in psychology* (pp. 197-215). Cambridge: Cambridge University Press.

Fontaine, J. R. J. (2005). Equivalence. In K. Kempf-Leonard (Ed.) *Encyclopedia of Social Measurement* (Vol. 1, pp. 803-813). New York: Academic Press.

Fontaine, J. R. J. (2008). Traditional and Multilevel Approaches in Cross-Cultural Research: An Integration of Methodological Frameworks. In F. J. R. Van de Vijver**,** D. A. Van Hemert**,** & Y. H. Poortinga (Eds.), *Multilevel Analysis of Individuals and Cultures* (pp. 65-92). New York: Lawrence Erlbaum Associates.

Fontaine, J. R. J. (2011). A fourfold conceptual framework for cultural and cross-cultural psychology: relativism, construct universalism, repertoire universalism, and absolutism. In F. J. R. van de Vijver, A. Chasiotis, & S. M. Breugelmans (Eds.), *Fundamental questions in Cross-Cultural Psychology* (pp. 165-189). Cambridge: Cambridge University Press.

Fontaine, J. R. J., & Fischer, R. (2011). Data analytic approaches for investigating isomorphism between the individual-level and the cultural-level internal structure. In D. Matsumoto & F. van de Vijver (Eds.), *Cross-cultural research methods in psychology* (pp. 273-298). Cambridge, MA: Cambridge University Press.

Fontaine, J. R. J., Luyten, P., de Boeck, P., Corveleyn, J., Fernandez, M., Herrera, D., & Ittzés, A. (2006). Untying the Gordian Knot of Guilt and Shame- The structure of guilt and shame reactions based on situation and person variation in Belgium, Hungary, and Peru. *Journal of Cross-Cultural Psychology*, *37*, 273 -292.

Fontaine, J. R. J., Poortinga, Y. H., Setiadi, B. & Suprapti, S. M. (2002). Cognitive structure of emotion terms in Indonesia and The Netherlands. *Cognition & Emotion, 16,* 61-86.

Fontaine, J. R. J., & Scherer, K. R. (2013). The global meaning structure of the emotion domain: Investigating the complementarity of multiple perspectives on meaning. In J. R. J. Fontaine, K. R. Scherer, & C. Soriano (Eds.), *Components of emotional meaning: A sourcebook* (pp. 106-125). Oxford, UK: Oxford University Press.

Fontaine, J. R. J., Scherer, K. R., & Soriano, C. (Eds.). (2013). *Components of emotional meaning: A sourcebook.* Oxford, UK: Oxford University Press.

Fontaine, J. R. J., Scherer, K. R., Roesch, E. B., & Ellsworth, P. (2007). The world of emotion is not two-dimensional. *Psychological Science*, *18*, 1050-1057.

Frijda, N. H. (1986). *The emotions*. Cambridge, UK: Cambridge University Press.

Frijda, N. H., Kuipers, P., & ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology*, *57*, 212-228.

Frijda, N. H., Markam, S., Sato, K., & Wiers, R. (1995). Emotions and emotion words. In J.A. Russell, J.-M. Fernández-Dols, A. S. R. Manstead, & J. C. Wellenkamp (Eds.), *Everyday Conceptions of Emotion: An Introduction to the Psychology, Anthropology and Linguistics of Emotion* (1st ed., pp. 121-144). Dordrecht: Kluwer Academic Publishers.

Jackson, J. C., Watts, J., Henry, T. R., List, J.-M., Forkel, R., Mucha, P. J., Greenhill, S., Gray, R. D., & Lindquist, K. A. (2019). Emotion semantics show both cultural variation and universal structure. *Science, 366,* 1517-1522.

James, W. (1884). What is an emotion? *Mind*, *9*, 188-205.

Kitayama, S., King, A., Yoon, C., Tompson, S., Huff, S., & Liberzon, I. (2014). The Dopamine D4 Receptor Gene (DRD4) Moderates Cultural Difference in Independent Versus Interdependent Social Orientation. *Psychological Science, 25*, 1169-1177.

Kitayama, S., Markus, H. R., & Kurokawa, M. (2000). Culture, emotion, and well-being: Good feelings in Japan and the United States. *Cognition & Emotion, 14*, 93-124.

Klineberg, O. (1940). *Social psychology.* New York: Henry Holt.

Kret, M. E., Stekelenburg, J. J., Roelofs, K., & de Gelder, B. (2013). Perception of face and body expressions using electromyography, pupillometry and gaze measures. *Frontiers in Psychology*, *4*, 28. https://doi.org/10.3389/fpsyg.2013.00028

Kuppens, P., Ceulemans, E., Timmerman, M. E., Diener, E., & Kim-Prieto, C. (2006). Universal intracultural and intercultural dimensions of the recalled frequency of emotional experience. *Journal of Cross-Cultural Psychology, 37*, 491-515.

Lang, P. J, Bradley, M. M. & Cuthberth, B. N. (1999). *International Affective Picture System (IAPS): Technical Manual and Affective ratings.* Gainesville, FL: University of Florida.

Lange, C. (1922). *The emotions* (I. A. Haupt, Trans.). Baltimore: Williams & Wilkins.

Lange, J., Dalege, J., Borsboom, D., Van Kleef, G. A., & Fischer, A. H. (in press). Toward an integrative psychometric model of emotions. *Perspectives on Psychological Science.*

Levenson, R. W., Ekman, P. E., Heider, K. & Friesen, W. V. (1992). Emotion and autonomic nervous system activity in the Minangkabau of West Sumatra. *Journal of Personality and Social Psychology, 62,* 972-988.

Loijens, L., & Krips, O. (2018). *Face Reader Methodology Note*. Noldus Information Technology White Paper.

Lutz, C. (1988). *Unnatural emotions: Everyday sentiments on a Micronesian atoll and their challenge to Western theory*. Chicago: University of Chicago Press.

Matsumoto, D., & Yoo, S. H. (2006). Toward a New Generation of Cross-Cultural Research. *Perspectives on Psychological Science, 1*, 234-250.

Manstead, A. S. R. & Fischer, A. H. (2002). Beyond the universality-specificity dichotomy. *Cognition & Emotion, 16*, 1-9.

Menon, U., & Shweder, R. A. (1994). Kali's tongue: Cultural psychology and the power of shame in Orissa, India. In S. Kitayama & H. R. Markus (Eds.), *Emotion and culture: Empirical studies of mutual influence* (pp. 241-282). Washington, DC, US: American Psychological Association.

Mesquita, B., & Frijda, N. H. (1992). Cultural variations in emotions: A review. *Psychological Bulletin*, *112*, 179-204.

Mesquita, B., Frijda, N. H., & Scherer, K. R. (1997). Culture and emotion. In J. W. Berry, P. R. Dasen, & T. S. Saraswathi (Eds.), *Handbook of Cross-Cultural Psychology, Volume 2: Basic Processes and Human Development* (pp. 255-297). Needham Heights, MA: Allyn & Bacon.

Messick, S. (1989). Validity. In R.L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York: Macmillan.

Osgood, C. E., May, W. H., & Miron, M. S. (1975). *Cross-cultural universals of affective meaning*. Urbana, IL: University of Illinois Press.

Poortinga, Y. H. (1997). Towards convergence? In J. W. Berry, Y. H. Poortinga and J. Pandey (eds.), *Handbook of cross-cultural psychology, Vol. 1, Theory and Method* (2nd ed., pp. 347-387). Boston: Allyn & Bacon.

Poortinga, Y. (2015). Is "Culture" a Workable Concept for (Cross-)Cultural Psychology?. *Online Readings in Psychology and Culture, 2*(1). https://doi.org/10.9707/2307-0919.1139

Poortinga, Y. H., van de Vijver, F. J. R., Joe, R. C. & van de Koppel, J. M. H. (1987). Peeling the onion called culture: A synopsis. In Kağitçibaşi, Ç. (Ed.). *Growth and progress in cross-cultural psychology* (pp. 22-34). Berwyn, PA, US: Swets North America.

Riegel, M., Moslehi, A., Michałowski, J. M., Żurawski, Ł., Horvat, M., Wypych, M., Jednoróg, K., & Marchewka, A. (2017). Nencki Affective Picture System: Cross-Cultural Study in Europe and Iran. *Frontiers in Psychology, 8,* 274. doi: 10.3389/fpsyg.2017.00274

Rosaldo, M. Z. (1980). *Knowledge and passion: Ilongot notions of self and social life*. Cambridge: Cambridge University Press.

Roseman, I. J. (1991). Appraisal determinants of discrete emotions. *Cognition & Emotion*, *5*, 161-200.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*, 1161-1178.

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of cross-cultural studies. *Psychological Bulletin, 115*, 102-141.

Sbarra, D. A. (2015). Divorce and Health: Current Trends and Future Directions. *Psychosomatic Medicine, 77*, 227-236.

Schachter, S., & Singer, J. E. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, *69*, 379-399.

Scherer, K. R. (1999). On the sequential nature of appraisal processes: Indirect evidence from a recognition task. *Cognition & Emotion*, *13*, 763-793.

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, *44*, 695-729.

Scherer, K. R. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition & Emotion*, *23*, 1307-1351.

Scherer, K. R., & Fontaine, J. R. J., & Soriano, C. (2013). CoreGRID and MiniGRID: Development and validation of two short versions of the GRID instrument. In J. R. J. Fontaine, K. R. Scherer, & C. Soriano (Eds.), *Components of emotional meaning: A sourcebook* (pp. 523-541). Oxford, United Kingdom: Oxford University Press.

Scherer, K. R., & Wallbott, H. G. (1994). Evidence for universality and cultural variation of differential emotion response patterning. *Journal of Personality and Social Psychology*, *66*, 310-328.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology, 25*, 1-65.

Schwartz, S. H. (2006). A theory of cultural value orientations: Explication and applications. *Comparative Sociology, 5,* 136-182.

Soudijn, K. A., Hutschemaekers, G. J. M., & Van De Vijver, F. R. J. (1990). Culture conceptualisations. In F. J. R., Van de Vijver, & Giel J. M. Hutschemaekers (Eds.), *The conceptualisation of culture* (pp. 19-39). Tilburg: Tilburg University Press.

Tangney, J. P. (1990). Assessing individual differences in proneness to shame and guilt: Development of the self-conscious affect and attribution inventory. *Journal of Personality and Social Psychology*, *59*, 102-111.

Thayer, R. E. (1989). *The biopsychology of mood and arousal.* New York: Oxford University Press.

Tomkins, S. S. (1962). *Affect, imagery, consciousness: Vol. 1. The positive affects*. New York: Springer.

Tomkins, S. S. (1963). *Affect, Imagery, Consciousness: Vol. 2. The negative affects*. New York: Springer.

Van Osch, Y. M. J., Zeelenberg, M., & Breugelmans, S. M. (2016). On the context dependence of emotion displays: Perceptions of gold medalists' expressions of pride. *Cognition and Emotion*, *30*, 1332-1343.

Van de Ven, N., Zeelenberg, M., & Pieters, R. (2009). Leveling up and down: The experiences of benign and malicious envy. *Emotion, 9*, 419-429.

Van de Vijver, F. J. R., & Leung, K. (1997). *Methods and data analysis of comparative research*. Thousand Oaks, CA: Sage.

Van de Vijver, F. J. R., & Leung, K. (2011). Equivalence and bias: A review of concepts, models, and data analytic procedures. In D. Matsumoto & F. van de Vijver (Eds.), *Cross-cultural research methods in psychology* (pp. 273-298). Cambridge: Cambridge University Press.

Van de Vijver, F. J. R., van Hemert, D., & Poortinga, Y. H. (2008). Conceptual issues in multilevel models. In F. J. R. Van de Vijver**,** D. A. Van Hemert**,** & Y. H. Poortinga (Eds.), *Multilevel Analysis of Individuals and Cultures* (pp. 65-92). New York: Lawrence Erlbaum Associates.

Van de Vijver, F. J. R., & Poortinga, Y. H. (2002) Internal structure equivalence in multilevel research. *Journal of Cross-Cultural Psychology,* 33, 141-156.

Van de Vijver, F. J. R., & Tanzer, N. K. (1997). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology-Revue Europeenne De Psychologie Appliquee, 47*, 263-280.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98*, 219-235.

Wong, Y., & Tsai, J. (2007). Cultural models of shame and guilt. In J. L. Tracy, R. W. Robins, & J. P. Tangney (Eds.), *The self-conscious emotions: Theory and research* (pp. 209-223). New York, NY: Guilford Press.

Wundt, W. (1896). *Grundriss der Psychologie.* Leipzig: Engelmann.

Yik, M. S. M., Russell, J. A., & Barrett, L. F. (1999). Structure of self-reported current affect: Integration and beyond. *Journal of Personality and Social Psychology*, *77*, 600-619.