

The Trials and Tribulations of Cross-Cultural Research

Michele J. Gelfand

Department of Psychology

University of Maryland

In press

The Trials and Tribulations of Cross-Cultural Research

Those of us who do cross-cultural research usually have a story of precisely when we began our journey into the global territory of research. For some, the interest might have developed at a very young age; for others it might have been happenstance, having stumbled into the field after doing work within one culture for many years. The litany of reasons why we venture into the world of cross-cultural research is also very diverse—for some it might be to discover general universal principles; for others it is to uncover the “thick description” of a particular culture (Geertz, 1973); and still for others it might be to simply understand the complex elephant of what culture is, addressing fundamental issues of human nature.

Although our beginnings and passions for doing cross-cultural research are varied, we ultimately share many common experiences along the bumpy road of cross-cultural research—experiencing many of the joys and delights on the one hand, and many of the trials, the frustrations, and disappointments on the other. Unfortunately, the restrictions of journal space and norms that focus us on reporting hypotheses, procedures, and results make it difficult to collectively realize that many of the issues that we encounter along the way are, in fact, common. The conference on *Conducting Multinational Research Projects in Organizational Psychology: Challenges and Opportunities* that took place at Michigan State in October of 2009 was a great success in helping us to share our research stories and to collectively make *explicit* what is often not discussed or implicit across many individuals. Whether you call it an academic conference detailing complex research programs or a collective therapy session, sharing these stories is critical for building institutional knowledge in cultural science, for giving a realistic preview to newcomers entering the field, and for empowering all of us with information that can help facilitate doing high quality cross-cultural research.

In this chapter, I detail some of my own experiences with conducting cross-cultural research. At the conference I discussed the details of several large-scale projects that I conducted, including a quantitative study of tightness-looseness across 33 nations (Gelfand et al., 2011) and a qualitative study of subjective culture in the Middle East that I embarked upon across 8 nations (Gelfand, 2008). In this chapter I draw on my experiences doing these projects and I broaden the discussion to include other studies I conducted as well. In what follows, I begin with my own serendipitous entry into the field, and place it in a historical context, discussing the intellectual heroes who have had a great influence on my own thinking. I then discuss some lessons learned, accumulated across many studies and from a lot of time spent in the cultural trenches. Because there are many good academic treatments of the issues that arise in the cross-cultural research process (see the volume on Methodology in the seminal *Handbook of Cross-Cultural Psychology*, Triandis, 1980; see also Cohen, 2007; Gelfand, Raver, Holcombe Ehrhart, 2002; Matsumoto & Van de Vijver, 2011), in this chapter I discuss specific examples of my own work, some published and some unpublished, to more vividly illustrate the issues that are invariably encountered when venturing into cross-cultural research territory.

Stumbling into Cross-Cultural Psychology

*We shall not cease from exploration
and the end of all our exploring will be to
arrive where we started... and know the
place for the first time.*

T.S. Eliott

My own personal journey into cross-cultural research began, as T.S. Eliot (1943) so aptly noted, when I was forced to step out of my cultural comfort zone and had the acute realization of

just how much that had been fundamentally shaped by American culture. I was a junior in college, a sheltered kid from Long Island, when I ventured off to London for a semester. I remember the “strange” sounds, sights, and smells of the U.K., and being completely overwhelmed—dazed and confused—from experiencing the “culture shock” that comes along with being away from one’s own familiar territory. I vividly remember a phone conversation with my father that was arguably the beginning of my journey of becoming a cross-cultural psychologist. I was telling him about how strange it was that people in my study abroad group would just go to Paris, Amsterdam, Scotland, and the like, for just for a few days. My father responded, in his Brooklyn accent, “*Well imagine it’s like going from New York to Pennsylvania!*” That metaphor gave me so much comfort that the very next day I booked a low budget tour to Egypt. It was just like going from New York to California, I reasoned (much to my father’s dismay!). Those travels, and later living on an Israeli Kibbutz, sparked a life long passion for understanding the dynamics of culture. I was fascinated with basic questions such as: How is it that culture shapes the self so profoundly yet culture is so invisible and taken for granted? How does culture develop, how is it sustained, and how does it change over time? How does culture contribute to misunderstandings and conflict at the individual, organizational, and national levels?

When I was back at Colgate University for my senior year, I was fortunate to find that Carolyn Keating, a cross-cultural psychologist, was teaching a cross-cultural human development class. Keating was a student of Marshall Segall, and it was in that course that I became exposed to their great work from the 1960s in Africa that showed that even basic psychological processes—such as visual perception—are not necessarily universal. At the time, the notion that humans might vary in fundamental ways of perceiving the physical world—space,

size, distance, or color—went largely unquestioned in psychology. Segall, Campbell, and Herskovits (1966) turned this assumption on its head. Taking a largely empiricist and Brunswikian perspective, they argued that people use whatever cues they have learned through their past experience to perceive objects, a process they referred to as *ecological cue validity*. Their research, done across 15 countries, indeed showed that Europeans were much more susceptible to classic illusions, such as the Müller-Lyer illusion, and the Sander Parallelogram illusion. Aside from these fascinating differences in deep psychological processes, I was intrigued by their explanation of the findings: In explaining such differences, Segall and colleagues advanced the *carpentered world hypothesis* which suggested that individuals who experience a lot of rectangular angles in their environment (which is more the case in Western cultures as compared to non-Western cultures) would be more likely to interpret non-rectangular figures as representations of rectangles, thereby exacerbating these types of visual illusions. I found this work—namely that *culture* is a prime source of experience that causes different habits of inference to arise—to be completely fascinating, and it was my first entrée into the wide world of cross-cultural research.

It was in that course that I also became inspired by work that had been done in the “culture and personality school” in the 1950s and 1960s, and in particular, on how socialization processes and personality factors vary *within* a particular society (i.e., Benedict, 1946; Mead, 1928), and how culture shapes, and is shaped by personality *across* different societies (Whiting & Child, 1953, Whiting & Whiting, 1975). I was intrigued by the ecological approach in Whiting and Whiting’s (1975) classic work *The Children of Six Cultures*. Whiting’s theory highlighted the role of the physical environment (e.g., climate, terrain), history (e.g., migrations), and maintenance systems (e.g., subsistence patterns, social structure), as important factors that shape

children's learning environment, which in turn, were thought to affect the development of adult personality (including learned and innate components) and projective expressive systems.

Although the theory had a largely deterministic flavor and drew heavily on psychoanalysis which was of less interest to me, the broad view of culture—and that idea that one could use the scientific method to understand it—was incredibly inspiring to that sheltered kid from Long Island!

After I graduated from college, I was determined to go to graduate school to study culture and psychology, yet it was clear that there were no Ph.D. degrees in this field. In a fateful conversation, Keating recommended that I talk to Richard Brislin, the then head of the East-West Center at the University of Hawaii and expert in cross-cultural training, about graduate programs in the field. After listening to my interests, Brislin declared that I work with Harry Triandis at Illinois. The rest was history.

Triandis was an incredible mentor who influenced my thinking and my approach to science. His sheer breadth and depth in the study of culture—both studying basic cultural processes and also culture's applications to personality, social, and organizational psychology—is forever inspiring to me. His early work on the *Analysis of Subjective Culture* (1972) is a classic that has influenced my thinking to this day. Triandis was highly influenced by the work of Melvin Herskovits (1955) who defined culture as the human-made part of the environment, consisting of both *physical* (e.g., tools, bridges, educational systems, religious institutions) as well as *subjective* elements (e.g., beliefs, attitudes, norms, values). Triandis set out to further explore “subjective culture” and was the first to develop methods to systematically identify social psychological constructs such as categorizations, associations, attitudes, beliefs, expectations, roles, and norms across cultural groups. His work showed that *coherent themes* cut

across these different elements of subjective culture. One such theme that was found in his seminal work was that of individualism and collectivism, which we later examined in terms of its vertical and horizontal elements (Triandis & Gelfand, 1998). Another major contribution of the *Analysis of Subjective Culture* (Triandis, 1972) was that, similar to Whiting (and later Berry, 1979), it placed the thematic elements of subjective culture into a larger ecological and historical framework. The theoretical framework that was developed included *Distal Antecedents* (e.g., climate) and *Historical Events* (e.g., wars), *Proximal Antecedents* (e.g., occupations, language used, religion), and *Immediate Antecedents of Action* (which included all the elements listed above), which result in patterns of action. The *Analysis of Subjective Culture* set the stage for large scale studies on dimensions of culture that took hold in the 1980s and had a strong influence on my later interest in tightness-looseness and the structure of everyday situations. Harry's work on the dynamics of culture (Trafimow, Triandis, & Goto, 1991) inspired my later interest in culture by situation interactions in the domain of negotiation and conflict (Gelfand & Realo, 1999; Morris & Gelfand, 2004).

Aside from the breadth of his theoretical training, Harry gave me a big dose of the methodological realities that one confronts in doing cross-cultural research. His admiration for both *emics* (culture-specific elements) and *etics* (culture-general elements), and his insistence on using multiple methods inspired me to use qualitative interviews, surveys, experiments, and archival methods, among others. He grounded me in the rich history and debates in the field, which I believe are critical to convey to new scholars (Kashima & Gelfand, 2011). Above all, Harry's optimism, modesty, and good humor helped to bring the human element into science. His philosophy—that it is important to be passionate about one's work, to not take yourself too

seriously, and to not be afraid to be controversial—have served as important reminders to me throughout my career on the bumpy road of doing cross-cultural research.

Views from the Trenches: Trials and Tribulations in Doing Cross-Cultural Research

As a graduate student in the early 1990s I read many papers about the issues that one confronts when doing cross-cultural research, but these challenges, which seem like distal abstractions as a student, come to life when you begin doing the work. The cross-cultural research process, as I described it (Gelfand et al., 2002), involves a roadmap wherein one has to make numerous “judgment calls” or crucial decisions that need to be made without a hard or fast rule (McGrath, 1982). The research process is also *cultural*, in that the very issues that we often seek to study—values, norms, beliefs, assumptions—also infiltrate every stage of the process, from deciding what is a worthy question (which itself is value laden) (Gelfand, Leslie, & Fehr, 2008), to creating a multicultural research team (in which cultural intelligence is critical for managing cultural diversity), to designing and implementing a particular research method (which involves unique cultural reactions, differential motivations, and ethical issues), to analyzing, interpreting, and publishing the data. In this respect, methods are infused with culture; they are difficult to separate. As Shweder (1990) once aptly remarked “you can't take the stuff out of the psyche and you can't take the psyche out of the stuff” (p.22). So too is the case with conducting cross-cultural research.

Theory as the starting point

Probably the best advice I got when starting to do cross-cultural research was the importance of having a strong theory. Because of the sheer number of ‘rival hypotheses’ that can explain one’s findings (other than culture), having a strong theory provides one with much greater confidence that the results are not due to extraneous factors. Theory here refers to both

the constructs of interest (what is the construct space? how is it operationalized?) and the relationships between constructs. Cross-cultural research is complex because the theory of the construct itself requires serious consideration, particularly given the risk of “imposing” etic constructs in other cultures (Berry, 1980). On the one hand, many, including myself, deal with this issue statistically—can one demonstrate that the scales being used have similar factor structures, similar loadings, and are in essence “equivalent” (Berry, Poortinga, Segall, and Dasen, 1992). One of my first research projects on sexual harassment did just this. I used simultaneous factor analysis in multiple samples to demonstrate that the sexual experiences questionnaire (SEQ), which had been developed in the United States had the same structure in Brazil (Gelfand, Fitzgerald, & Drasgow, 1995). Over time, however, I became skeptical with using this statistical approach to justify the *theory* of the construct because it assumes that the construct space itself—which has almost always been derived from Western theory and Western samples in the vast majority of cases—has been “mapped” appropriately. From a psychometric perspective, while factor analysis can illustrate if certain items or dimensions are relevant (are similar) and/or are contaminated (are not working well in other cultures), it cannot reveal whether important dimensions of a construct have been neglected or omitted (are there certain dimensions of harassment that are novel in Brazil?). This is a theoretical issue and requires a deep understanding of the cultural context. I often find myself asking, why would a construct developed in the U.S. *necessarily* have the full range of variation that is needed to capture realities of the phenomenon in another context? For example, Ramesh and Gelfand (2010) showed that while job embeddedness was a universal predictor of turnover, family embeddedness was an important (but heretofore neglected) predictor of turnover in India (and even in the U.S). Fahr and Earley (1997) showed that if one wants to understand organizational

citizenship behavior (OCBs) in China, new dimensions (unearthed in interviews about what constitutes the construct) and new measures need to be added to the construct for it to be relevant in China. To be clear, it is not necessarily the case that new dimensions and items need to be added to a construct and/or measure; but assuming that the construct space ‘travels’ perfectly is problematic.

Theory also guides the choice of samples. For example, in early work that I did on culture and procedural justice, I theorized that voice was much more important in high power distance than low power distance cultures. I selected Turkey (which is high on power distance) and Costa Rica (which is low on power distance) because they are both collectivistic societies, thus trying to, generally speaking, isolate the cultural variable of interest. Our sampling of cultures for a multi-nation study on cultural tightness-looseness was likewise guided by theory of the predictors of the construct. I theorized that population density, history of territorial conflict, resource scarcity, human disease, and natural disasters are predictive of tightness-looseness, and consulted extant archival databases (see below discussion of the promise and pitfalls of archival databases) in order to choose nations that reflected substantial variation on these variables (Gelfand et al., 2011). In other work (Kashima, Yamaguchi, Kim, Choi, Gelfand, & Yuki, 1995), we sampled nations so as to have variability on the individual, collective, and relational self in order to examine culture and gender influences on these constructs. As compared to representative sampling, this approach reflects *theoretical* sampling (Boehnke, Lietz, Schreier, & Wilhelm, 2011) that aims to maximize the variability of nations according to the theory being tested.

Theory also guides measures, whether they are mediators, moderators, or control variables. The importance of “unpacking” cultural differences has long been discussed in cross-

cultural psychology (Whiting & Whiting, 1975). Merely showing country or other group differences does not elucidate the *reasons* for the effects, making it critical to try to illuminate the cultural phenomenon explaining differences found. Although much work has used survey measures of personal values (for example, of individualism-collectivism) as potential mediators, cultural mediators need not be based on value measures on surveys—they can be based on descriptive norms (Shytenberg, Gelfand, & Kim, 2008; Zou et al., 2008), the structure of situations, roles, and networks (Kitayama, Markus, Matsumoto, & Norasakkunkit, 1997; Gelfand et al., 2011; Morris, Poldony, & Ariel, 2000), implicit measures (Kitayama & Karasawa, 1997), and artifacts and/or cultural products (Morling & Leumeax, 2008), among other variables. Unpacking cultural differences is a tricky business; but ultimately theory is critical for identifying potential mechanisms and alternative explanations for country or group differences.

Cultural legwork

In my experience, compared to doing research on one sample, cross-cultural research takes much longer and requires considerably more “leg work” prior to even launching a study. For example, gaining access to samples and developing a highly functioning cross-cultural research team is critical for a successful cross-cultural project and requires time, resources, cultural skills, a lot of patience, and some degree of pure luck. Building networks of people who are interested in the same questions, committed to the same publication goals, are willing and able to do all that is required to complete the process (translations, piloting, implementing, and interpreting data), and who are on the same time frame with the necessary resources, is a daunting task. Managing expectations—which can be very different depending on one’s cultural background—is critical from the very start of research projects. Basic questions that need to be addressed prior to the launching of a project include: Will all collaborators be authors on the

study? (I personally believe this is important, having been influenced on this issue with Triandis). Is the research question and method appropriate in the local context? How will translations be handled? Who “owns the data?” Are there funding agencies involved, and does this have implications for local collaborators? Are there ethical issues that need to be discussed involving the participants and/or the researchers? For example, in a project that I have been involved in examining culture and negotiation in the Middle East, the funding agency (Department of Defense) clearly wants to be acknowledged on publications that come from the project. Yet for some projects, it became clear that acknowledging the DOD would place collaborators in some countries at a huge reputational risk. Knowing this, I was able to negotiate with the DOD that in some cases we cannot acknowledge the funding agency. Having honest discussions about the goals of the research, where one wants to publish it, authorship, funding, and timeframes for the work are critical for the success of a multicultural research team. Of course, “honest discussions” themselves are culturally constructed, and considerable trust and cultural skills are critical for dealing with controversial issues, and the invariable relationship, task, and process conflict that can occur in multicultural collaborations. There is the added temptation, in my experience, of dealing with these issues using email discussions, which themselves are too “lean of a medium” (Daft & Lengal, 1984) to handle the complexity of these tasks. In my own work, I try to meet often face to face, talk on Skype (at least you can gauge non-verbals) and communicate frequently in order ensure that everyone is fully comfortable with the many issues involved in the research.

Having agreed upon the nature of the collaboration, one then is in a position to get the research started. This itself also takes considerable time, energy and patience. Whereas in my own uni-cultural work in the lab or field, I can typically get my surveys or experiments ready for

testing relatively quickly, implementing cross-cultural research involves many other steps. When using an experiment or survey across cultures, it is critical to discuss and modify the questions, methods, and designs with local collaborators, and critically to *expect this cultural input to change one's plans*.

For example, in a recent collaboration, I created a multicultural team involving collaborators from Egypt, Iraq, Jordan, Lebanon, Pakistan, Turkey, and the United Arab Emirates in order to understand important subjective culture—values, focal concerns, beliefs, and norms—in the Middle East (ME) region. Because there is very little research on the psychology of culture in the Middle East, I thought it was critical to use qualitative methods to gain a “thick description” (Geertz, 1973) prior to developing surveys, experiments, or other methods. I settled on conducting interviews in all of these countries using the methodology pioneered in the *Analysis of Subjective Culture* (Triandis, 1972) as the platform, which involves word associations, antecedents and consequences of relevant constructs, and methods to illuminate situational variation in the constructs. Prior to developing the interview protocols, I read through numerous books and articles about the region, and then drafted a list of potential constructs to discuss with the ME team. These constructs were then the subject of many discussions within the team, and various questions were dropped as other questions were added. Based on this input, I drafted the next version of the interview protocols, which included interview probes on the psychology of connections (*wasta*), fate, honor (*sharaf, irdh*), face and public image (*wujah*), respect, modesty, dignity, values, trust, negotiation, conflict, revenge, forgiveness and apologies, and collaboration. This next draft was then the subject again of numerous discussions and iterations. After several months, we settled on a final set of protocols to pilot with local focus groups. Although the process was very time consuming, with this

extensive cultural legwork, I could be much more confident that the interview protocols reflected important focal concerns and would yield valuable “cultural fruit.”

Prior to even piloting the protocols (which is critical to do before launching the interviews), we had a lot more work to do to get ready to launch the study. First, there was the anguishing process of translations. All protocols were originally written in English and had to be translated into Arabic, Farsi, Turkish, and Urdu. Once translated into the local language, we had different translators backtranslate them into English so we could compare the original English version and backtranslated version. Invariably there were many discrepancies, so we then needed to figure out where the difference in translation occurred (was it a problem in the original translation? Or was the backtranslation the culprit?). In my experience, translations often reveal problems that have a basis in *poor English*, for example, double barreled statements, colloquial language, metaphors and idioms that don't even make sense to English speakers! As a rule of thumb, I have always assumed that the English version will need to be modified in addition to the translated version. Another rule of thumb is that the one should estimate the time needed to do translations and then double or even triple this time to have a realistic estimate! Coordinating numerous translations (wherein there might need to be different changes that are standardized across all materials) takes considerable time. For example, in our study of tightness-looseness across 33 nations, our study materials were in 21 languages involved, and the translation process took fully one year to complete.

When the materials are translated and one has confidence with the final materials, it is time to pilot the methods. Piloting is critical for the quality of cross-cultural research projects (Gelfand et al., 2002). All methods—experiments, interviews, surveys—need to be carefully examined prior to the implementation of research for numerous criteria: Are participants familiar

with the tasks, comfortable with them, motivated by them? Do they understand the instructions in the intended way? How do they react to the experimenter, interviewer or otherwise “epistemic authority” running the study? Are they perceived as ethically acceptable? Is there a problem with using deception if applicable? Is the ample time for the study? Are the incentives appropriate in all cultural groups? Piloting the study prior to data collection, in my experience, always results in important changes. For example, our pilots of the interview study discussed above revealed a lot of perceived overlap in the questions that made participants less motivated. As a result of this feedback, we had to further condense and re-organize the protocols. Our pilots also revealed that we needed to spend more time building rapport, particularly in rural areas, in order to make people more comfortable answering the questions, some of which were highly sensitive.

Many of the pilots that I have conducted illustrate how the methodological choices we make are often laden with Western assumptions, ironically that can relate to the very questions we are asking. Put differently, the very “stuff” that we’re interested in looking at cross-culturally is found in cultural issues that one encounters in the method. For example, in Gelfand and Realo (1999) we were interested in examining how accountability produces very different effects in negotiation depending on the cultural context. Much research on accountability in the U.S. in the 1970s and 1980s showed that accountability produces competition in negotiation: In fact, negotiators assume that their constituencies want them to be competitive (Benton & Druckman, 1973; Gruder, 1971). Not surprisingly, accountability activated competitive construals and behaviors, and resulted in lower negotiation outcomes for individualistic samples. We reasoned that accountability need not produce competition in all cultures as it does in the individualistic U.S.; rather, we hypothesized that accountability acts as a norms enforcement mechanism, producing whatever is normative in the cultural environment. To test this, we first conducted

research in the U.S. among Asian Americans and Caucasians and showed that in fact among collectivists (Asian participants), accountability activated cooperative construals and behaviors, and resulted in higher negotiation outcomes. These effects were reversed in unaccountable negotiations, when in effect, negotiators were released from normative pressures to do what is expected. Interestingly, when we tried to pilot the material in Japan for a follow up study, we had a very difficult time getting our manipulations of accountability to work. In fact, low accountability situations were seen as having very high accountability, even after numerous pilots, making it difficult for us to run the study in Japan (at least as a scenario study). In retrospect, getting the manipulations of accountability to be equivalent in the U.S. and Japan is not particularly surprising given that Japan is a very tight culture with much higher monitoring (Gelfand et al, 2011) and thus has higher naturally-occurring accountability as compared to the U.S. Thus, what is perceived to be a low accountability situation in one culture may be seen as a high accountability situation in a culture with a higher base rate, in general, of monitoring. This one example aside, it is critical to ensure that one's manipulations are understood in the same way and have equal strength prior to data collection, even if the manipulation has worked time and time again in one's own cultural context.

In addition to having equal *cognitive comprehension* of the task instructions, it is important to ensure that participants in all cultures have *equal motivation* to perform the task. In other words, comprehension of the task is not enough; it also must be equally engaging across groups. Again, ironically enough, the very nature of the research question that is being addressed can reveal cultural issues in motivation in the method. For example, in Gelfand, Higgins, Raver, Nishii, Dominguez et al. (2002) we were interested in cross-cultural differences in egocentric biases in negotiation. Much research in the field of negotiation has illustrated that negotiators

tend to view their own behaviors as more fair than others (Thompson & Loewenstein, 1992), which leads to more aggressive behavior, less concession-making, and ultimately lower outcomes (Babcock & Loewenstein, 1997). In Gelfand et al. (2002), we theorized that serving biases of fairness in negotiation would be consistent with ideals within individualistic cultures, where the self is served by focusing on one's positive attributes in order to "stand out" and be better than others, but would be disruptive to ideals in collectivistic cultures, where the self is served by focusing on one's negative characteristics in order to "blend in" and maintain interdependence with others (Heine, Lehman, Markus, & Kitayama, 1999). We first did a number of survey and scenario studies to examine cultural differences in egocentric biases of fairness in conflict in the U.S. and Japan, and then in our last study, we set out to examine how differences in egocentric biases affect "hard" negotiation outcomes. We created a simulation that required students to assume rules and to negotiate over four issues. The context for the negotiation, we thought, was very interesting and engaging—namely a negotiation between two "distinguished" honor student clubs over space, time, and other issues on which they needed to coordinate. We piloted the task in the U.S., and it was a smashing success. Students reported it was very motivating and enjoyable. When piloting in Japan, however, we found that the task completely *flopped*—the students reported being uncomfortable and unmotivated with the issues that in their cultural eyes seemed very atypical. It became clear from discussions that this was a "vertical individualistic task" and produced negative reactions in Japan. Unbelievably, the very same question that we were interested in (cultural differences in how much people like to stand out and its implications for negotiation in different cultures) became embedded in our research tasks! We went back to the drawing board and developed a new buyer and seller task that proved to be much more motivating and relevant in both contexts (Gelfand et al., 2002). Without this

extensive piloting, we would have missed the fact that even though it would have been possible to translate the original honor's club task, and have it be equally comprehensible in both cultures, participants would not have been equally motivated to engage in the simulation in Japan and our results would be very difficult to interpret.

In other work we have done, pilots and focus groups often reveal that the methods that we export abroad are far too *decontextualized* to be understandable and motivating in other cultures. In our interview questions, for example, we initially asked questions such as “Is compromise good or bad?” and later adapted the questions to have tags to reflect the *circumstances* under which is it good or bad. In our interviews, respondents in general in the Middle East needed to know, with whom, about what issues, and in what circumstances we were asking about compromise, negotiation, *wasta*, revenge, forgiveness, apologies, among other constructs. We rarely found that our American counterparts asked for this information. In experiments we have likewise found it important to include much more information in the case materials in order for the instructions and study materials to make any sense. For example, in our case *At Your Service* (Brett & Gelfand, 2004), we had individuals negotiate either as part of a team or as part of a dyad in the U.S. and Taiwan. The case involved a relationship between two individuals who were either trying to form a deal to own a restaurant together (i.e., a deal-making context) or were trying to dissolve a relationship that had been subject to many problems (i.e., a disputing context). When we first wrote the case, we included only details about the issues to be resolved. Yet pilot studies in Taiwan revealed that they couldn't negotiate the case without more information about the people, their histories, their relationship, etc. Put simply, it was too “low context” for a high context culture (Hall, 1976).

Other Western assumptions also sneak into our methods. For example, in our “time as money” Western culture, we often assume that participants can read the materials, get into roles, answer questions, and negotiate a case in a remarkably short time. In the same Taiwan-U.S. study discussed above, we ran the study in 90 minutes, but we found that over 50% of the Taiwanese participants failed to reach agreement in that time. We found some very interesting results, but in order to trust the findings, we had to re-run the entire study to ensure that our results were not a methodological artifact of not having enough time to negotiate. Ultimately, the results of our follow up experiment were identical, giving us confidence in the theory. Nevertheless, in this case, it was clear to us that the same amount of time that is needed in one culture to complete a study might be highly problematic in another.

In sum, the above descriptions and examples make clear that it is important to choose a task collaboratively with all local researchers that will ensure equal familiarity, comprehension, motivation, and ethicality in the study and illuminate any potential problems in the implementation of the task. Either focus groups or pilot analyses should be held in each culture, and the results of these preliminary analyses should be used to make substantive changes in the protocols. In the pilot test or focus groups, I have found it useful to employ comprehension and motivation checks (Berry, 1980), or use judgmental methods with which experts evaluate the stimuli (Segal et al., 1992).

Methodological Tradeoffs and the Importance of Triangulation

McGrath (1982) always instilled in me the idea that all methods are flawed and each has strengths and weaknesses. In cross-cultural research, all methods have additional cultural baggage (Gelfand et al., 2002), and many rival hypotheses can threaten one’s confidence in the interpretation of cultural differences that are found. Because of this, I have always found that to

the greatest extent possible, it is important to see one's theory replicated with more than one method—that is, to see that the results triangulate.

For example, *questionnaires or surveys* have a number of advantages including that they may be less intrusive than other methods (e.g., laboratory experiments, discussed below) and provide the ability to collect data on a wide range of questions at any one time. Cross-cultural challenges to surveys abound, however, including potential differences in motivation, understanding of instructions, validity, reliability, and response sets, making it important to replicate findings with another method. In our study of cultural tightness-looseness (Gelfand et al., 2011), we measured the strength of norms and degree of sanctioning across 33 nations with Likert survey measures (e.g., “There are many social norms that people are supposed to abide by in this country,” “In this country, if someone acts in an inappropriate way, others will strongly disapprove,” “People in this country almost always comply with social norms,” among other items). Because survey measures can be subject to response sets and a lack of equivalence, it is important to perform procrustes factor analysis in all cultures to examine the structure of the measure and also check to see if standardization is required. In addition to these issues, differential motivation to respond (e.g., social desirability) and potential differences in the interpretation of the items also make it critical to gather additional data from other sources in order to provide convergent validity for one's survey measure. For example, we have more confidence in the Likert scale given that scores on our tightness-looseness scores were strongly correlated with “non-survey” measures including expert ratings on the construct, higher monitoring in society (more police per capita), and more severe punishments (e.g., the death penalty) for crime. The questionnaire measure was also correlated with unobtrusive measures including the percentage of people who write with their left hand (a very visible indication of being “deviant” from norms), and with

greater accuracy of public clocks in cities, indicating a greater concern with order and uniformity. By showing that our tightness-looseness scores were correlated with other indicators, rival hypotheses due to response sets or differences in meaning of items, among many other issues, were reduced. In another study that relied upon surveys of perceptions of conflict episodes in which we used multidimensional scaling to examine the dimensions on which people perceive conflicts in the U.S. and Japan (Gelfand et al., 2001), we coupled this method with analyses of newspaper accounts of conflicts in the New York Times and Japanese Yamiuri. Both analyses illustrated that while cooperation versus competition (or win-lose) frames were universal, U.S. conflicts were perceived to be much more competitive (win oriented) than those in Japan. In this case, cultural documents such as newspapers can help to provide additional confidence that one's theory generalizes beyond one method.

As another example, *experiments* provide numerous advantages in that they provide a controlled research environment and allow for greater inferences in causal relationships. Laboratory research is also beneficial in that it enables one to assess implicit attitudes in addition to explicit self-reported attitudes. Yet laboratory research, particularly in cross-cultural settings, can present many problems that can reduce confidence in the interpretability of the results. As noted above, it is critical that laboratory tasks and procedures are equally understood and motivating to individuals across different cultures. The very artificiality of the laboratory environment (and role playing manipulations that I often use) can be unfamiliar to people outside of the West, making it critical to replicate one's theories with another method. In our research on accountability using laboratory experiments (Gelfand & Realo, 1999) or self-serving biases (Gelfand et al., 2002), for example, it was important to see that our effects were also replicated using scenario-based measures.

Interviews and other qualitative methods are another useful method in cross-cultural research in that they enable one to gain depth on a research question and often afford more of an understanding of “emic” or culture-specific perspectives that are especially useful in early stages of cross-cultural research. Interviews are also essential when dealing with illiterate populations. Many difficulties and judgment calls arise, however when implementing this method. For example, characteristics of the interviewer need to be carefully chosen for their cultural appropriateness (e.g., in the Middle East, using female interviewers for female samples). The ways in which interviewers gain trust, develop rapport, and “probe” the participants to answer questions can be very different in different cultures making it important to try to “negotiate” a standardized interview process with one’s collaborators prior to data collection. For example, in some cultures, revealing information about oneself is critical for the development of trust, yet in other cultures this would be seen as inappropriate and a practice is perceived as threatening the “objectivity” of the data collection. Prior to our interviews in the U.S. and Middle East, we searched the literature for “best practices” in interviewing (unsurprisingly, much of this was derived from interviews with Western samples), and with input from our collaborators, we designed an interview manual which was discussed, revised, and finalized based on the entire teams’ input. Other aspects of the interview process, such as tape or video recording, can have very different implications depending on the culture. For example, recording of interviews could be highly sensitive in some cultures—particularly in very tight cultures where there is a concern about authorities hearing one’s responses (e.g., our experience in Iran), making additional safeguards and assurances necessary. For example, in recent work on community negotiations in the U.S. and Egypt, we were not able to videotape the negotiations, as this would have been considered too invasive among Egyptians.

Once interview data has been collected, there are many issues that arise in determining how to extract, code, and interpret the data. For example, in our interview project, we needed to first transcribe all interviews (which were between one and two hours per participant) from their native language from audio files to actual text in Arabic, Turkish, Urdu, and English. Each interview produced on average of 15-20 pages of actual text. Because we wanted to code the data in a standardized way, we needed to develop a method to reliably extract the answers. It took approximately 6 months to develop a process in which we felt confidence. We first developed a standardized manual for all team members to discuss. For each interview question, the Middle East (ME) and US teams first completed extractions (i.e., the answers to questions) on 2-3 designated transcripts for reliability purposes. We took the most difficult interviews (those with the most variance in terms of where the answers could be found, some being found right after the questions were asked and others that found later in the interview as well). We computed reliability across U.S. and ME collaborators, and after several iterations and resolving several disagreements (and updating our manual), we were able to have trust in the reliability of the process.

This process was far from completed. After all the data was extracted (approximately one year later for all questions), we needed to figure out what to do with this extracted data! We again sought to develop a standardized code development process that all collaborators could agree upon and that was reliable and valid. We were very concerned at this stage with not imposing our Western glasses on the data and allowing *emic* (culture-specific) concepts and themes to be identified in each country. The process that we created ultimately involved three phases. In the first phase, ME and US teams separately examined the extracted answers for a particular interview question and constructed a list of possible codes or themes for their

respective countries. For the construct of honor, for example, 1,769 codes, or an average of 103 codes per question, were generated across the teams! The second phase involved sorting and organizing these codes at a conceptual level by the US team with input from our ME collaborators as well as with input from extant research. Finally, the third phase involved writing a coding manual that describe the code categories in detail and set forth procedures and guidelines for coding. Coding then is done with bi-lingual individuals with reliability checks. This process is being implemented for all questions in the interview protocols. Not surprisingly, these steps take a lot of time and resources. And as with other methods, it is important to triangulate interview findings with other methods to gain confidence in the results. For example, much of our interview analyses illustrated that honor and face loss becomes much more contagious across individuals in the Middle East than the U.S, generally speaking. We are currently in the process of replicating these results with free recall methods, laboratory experiments, and computational models.

Finally, *archival databases* also have notable strengths in cross-cultural research in that they provide another unobtrusive source of cross-cultural data. Examples include ethnographies, which provide in-depth information about a given culture, and cross-cultural databases on ecological, sociological, economic or political variables. These sources, however, also have notable weaknesses. Preexisting databases may only be available for a limited number of countries or variables. In addition, databases may label or assess constructs differently across cultures, and as such, comparisons are problematic. In addition, without a developed theory, the use of such sources can result in “dustbowl” empiricism. I have found it important to try to find convergent evidence when possible to help bolster my confidence in my own research using archival methods. For example, in the study of tightness-looseness, I wanted to test the notion

that ecological threats (e.g., lack of natural resources) and man-made threats (e.g., population density, human disease) are related to tightness. In order to assess the extent to which nations are subject to resource scarcity, I collected data on arable land, food production, food supply, and food deprivation from the Food and Agriculture Organization of the United Nations, as well as the percentage of farmland and access to safe water in each country from Kurian's world ranking. Data on population density were taken from the *Atlas of World Population History* in the year 1500 but also from the United Nations in 2000. To assess threats due to human disease, I located a number of different sources to triangulate the results. The index of historical prevalence of pathogens was taken from Murray and Schaller's (2010) research, in which they constructed the disease prevalence index based on early epidemiological atlases. The World Health Organization (WHO) provided data for years of life lost to communicable disease and prevalence of tuberculosis—a highly communicable disease. Mortality rates for children under 5 were also gathered from the United Nations. By finding the theoretically expected results with multiple sources, one can have more confidence in the results of archival analyses. In addition, as with other methods, it is comforting to replicate one's results with a completely different method. For example, in recent research we have been “priming” ecological conditions (such as population density) in the laboratory in order to examine our theory of tightness-looseness.

In sum, all methods clearly have strengths and weaknesses, and all are very useful, particularly in combination, when doing cross-cultural research. Each method varies considerably on its capacity for gaining depth about the phenomenon (interviews), control and causal inferences (laboratory experiments), unobtrusive non-reactive measures (observations, content analyses), and the ability to have standardized, structured responses (surveys). Above

all, each method presents problems for interpreting results across cultures, rendering it essential to replicate with a complementary method when possible.

Concluding Remarks

As this brief essay I hope has illustrated, cross-cultural research to me is a passion and is a lifelong journey that brings with it many joys. It also presents many challenges at all stages of the research process that require many difficult judgment calls that do not necessarily have one right or wrong solution. To cope with the long and bumpy road, I have always relied upon the use of theory, much cultural legwork and the wisdom of my collaborators, the triangulation of methods, and patience. Textbooks and journal articles on methodological issues in cross-cultural research can also provide very useful technical advice. It is also instructive, however, to share the stories, dilemmas, and serendipity that are behind the scenes of our research lives in order to more vividly illustrate the issues that are invariably encountered when venturing into cross-cultural research territory. This is both empowering, as many of us experienced at the Michigan State conference on *Conducting Multinational Research Projects in Organizational Psychology: Challenges and Opportunities*, and it is also critical for building institutional knowledge in an ever more growing field.

References

- Babcock, L., & Loewenstein, G. (1997). Explaining bargaining impasse: The role of self-serving biases. *Journal of Economic Perspectives*, 11, 109-125.
- Benedict, R. (1946). *The chrysanthemum and the sword: Patterns of Japanese culture*. Boston: Houghton Mifflin.
- Benton, A. A., & Druckman, D. (1973). Salient solutions and the bargaining behavior of representatives and nonrepresentatives. *International Journal of Group Tensions*, 3, 28-39.
- Berry, J.W. (1979). A cultural ecology of social behaviour. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology*, Vol. 12, (pp. 177-206). New York: Academic Press.
- Berry, J. W. (1980). Introduction to methodology. In H. C. Triandis and J.W. Berry (Eds). *Handbook of Cross-Cultural Psychology* (Volume 2, pp. 1-28). Boston: Allyn and Bacon.
- Berry, J.W., Poortinga, Y.H., Segall, M.H. and Dasen, P.R. (1992). *Cross-cultural psychology: Research and Applications*. New York: Cambridge University Press.
- Boehnke, Lietz, Schreier, & Wilhelm (2011). Sampling: The selection of cases for culturally comparative psychological research. In D. Matsumoto & F. Van de Vijver (Eds.) *Cross-Cultural Research Methods in Psychology* (pp. 101-129). Cambridge UK: Cambridge University Press.
- Brett & Gelfand (2008). At your service. *Dispute Resolution Research Center, Northwestern University*.
- Cohen, D. (2007). Methods in cultural psychology. In S. Kitayama & D. Cohen (Eds). *Handbook of Cultural Psychology* (pp 196-236). New York: Guilford Press.

- Daft, R. L. and Lengel, R. H. (1984). Information richness: a new approach to managerial behavior and organizational design. *Research in Organizational Behavior* 6, 191-233.
- Eliot, T.S. (1943/1971) *Four Quartets*. New York: Harcourt, Inc.
- Farh, J-L., & Earley, P.C. (1997). Impetus for action: A cultural analysis of justice and organizational citizenship behavior in Chinese society. *Administrative Sciences Quarterly*, 42, 421-444.
- Geertz, C. (1973). *The interpretation of cultures*. New York: Basic Books.
- Gelfand, M. J. (2008). Dynamical models of culture and collaboration in the Middle East. PI on Grant funded by the U. S. Army Research Laboratory and the U. S. Army Research Office (W911NF-08-1-0144).
- Gelfand, M. J., Fitzgerald, L. & Drasgow, F. (1995). The structure of sexual harassment: A confirmatory analysis across cultures and settings. *Journal of Vocational Behavior*, 47, 64-177.
- Gelfand, M. J., Higgins, M., Nishii, L., Raver, J., Dominguez, A., Yamaguchi, S., Murakami, F. & Toyama, M. (2002). Culture and egocentric biases of fairness in conflict and negotiation. *Journal of Applied Psychology*, 87, 833-845.
- Gelfand, M. J., Leslie, L., & Fehr, R. (2008). In order to prosper, organizational psychology...should adopt a global perspective. *Journal of Organizational Behavior*, 29, 493-517.
- Gelfand, M. J., Nishii, L. H., Holcombe, K., Dyer, N., Ohbuchi, K., & Fukumo, M. (2001). Cultural influences on cognitive representations of conflict: Interpretations of conflict episodes in the U.S. and Japan. *Journal of Applied Psychology*, 86, 1059-1074.

- Gelfand, M. J., Raver, J. L., & Holcombe Ehrhart, K. (2002). Methodological issues in cross-cultural organizational research. In S. Rogelberg (Ed.) *Handbook of industrial and organizational psychology research methods* (pp. 216-241). New York, NY: Blackwell.
- Gelfand, et al. (2011). Differences between tight and loose societies: A 33-nation study. *Science*, 332, 1100-1104.
- Gelfand, M. J. & Realo, A. (1999). Individualism-collectivism and accountability in intergroup negotiations. *Journal of Applied Psychology*, 84, 721-736.
- Gruder, C. L. (1971). Relationship with opponent and partner in mixed-motive bargaining. *Journal of Conflict Resolution*, 15, 403-416.
- Hall, E. T. (1976). *Beyond culture*. New York: Anchor Books.
- Heine, S. J., Lehman, D. R., Markus, H. R., & Kitayama, S. (1999). Is there a universal need for positive self-regard? *Psychological Review*, 106, 766-794
- Herskovits, M. J. (1955). *Cultural anthropology*. New York: Knopf
- Kashima, Y., & Gelfand, M. J. (2011). Culture: A brief history of meaning in psychology. In W. Stroebe & A. Kruglanski (Eds.), *History of Social Psychology*.
- Kashima, Y., Yamaguchi, S., Kim, U., Choi, S., Gelfand, M. J., & Yuki, M. (1995). Culture, gender, and self: A perspective from individualism-collectivism research. *Journal of Personality and Social Psychology*, 69, 925-937
- Kitayama, S., Markus, H., Matsumoto, H., & Norasakkunkit, V. (1997). Individual and collective processes in the construction of the self: Self-enhancement in the United States and Self-Criticism in Japan. *Journal of Personality and Social Psychology*, 72, 1645-1267.
- Matsumoto, D., & Van de Vijver, F. J. R. (2011). *Cross-cultural research methods in psychology*. Cambridge UK: Cambridge University Press

- McGrath, J. E. (1982). Dilemmatics: The study of research choices and dilemmas. In J. E. McGrath, J. Martin, and R. A. Kulka (Eds) *Judgment Calls in Research* (pp. 69-102). Beverly hills, CA: Sage.
- Mead, M. (1928). *Coming of age in Samoa*. New York: Morrow.
- Morling, B., & Lamoreaux, M. (2008). Measuring culture outside the head: A meta-analysis of individualism-collectivism in cultural products. *Personality and Social Psychology Review*, 12, 199-221.
- Morris, M. W. & Gelfand, M. J. (2004). Cultural differences and cognitive dynamics: Expanding the cognitive tradition in negotiation. In M. J. Gelfand & J. M. Brett (Eds.) *The handbook of negotiation and culture* (pp. 45-70). Palo Alto, CA: Stanford University Press.
- Morris, M., Podolny, J., & Ariel, S. (2000). Missing relations: Incorporating relational constructs into models of culture. In *Innovations in International and Cross Cultural Management*, 52-90. Ed. P. C. Earley and H. Singh. Thousand Oaks, Calif.: SAGE, 2000.
- Murray, D. R., & Schaller, M. (2010). Historical prevalence of infectious diseases within 230 geopolitical regions: A tool for investigating origins of culture. *Journal of Cross-Cultural Psychology*, 41, 99-108.
- Ramesh, A., & Gelfand, M. J. (2010). Should they stay or should they go: Job embeddedness in predicting turnover in individualistic and collectivistic cultures. *Journal of Applied Psychology*. 95, 807-823.
- Segall, M. H., Campbell, D. T., Herskovits, M. J. (1966). *The influence of culture on visual perceptions*. Indianapolis: Bobbs-Merrill.

- Shteynberg, G., Gelfand, M. J., & Kim, (2009). Peering into the ‘magnum mysterium’ of culture: The explanatory power of descriptive norms. *Journal of Cross-Cultural Psychology*.
- Shweder, R. A. (1990). Cultural psychology: What is it? In J. W. Stilger, R. A. Shweder, & G. Herdt (Eds). *Cultural psychology: Essays on comparative human development* (pp. 1-43). Cambridge, MA: Cambridge University Press.
- Thompson, L., & Loewenstein, G. (1992). Egocentric interpretations of fairness and interpersonal conflict. *Organizational Behavior and Human Decision Processes*, 51, 176-197.
- Trafimow, D., Triandis, H. C., & Goto, S. G. (1991). Some tests of the distinction between the private and the collective self. *Journal of Personality and Social Psychology*, 60, 649-655.
- Triandis, H.C. (1972). *The analysis of subjective culture*. New York: Wiley.
- Triandis, H. C. (1980). *The Handbook of Cross-Cultural Psychology*. XXX.
- Triandis, H.C., & Gelfand, M.J. (1998) Converging measurement of horizontal and vertical individualism and collectivism. *Journal of Personality and Social Psychology*, 74, 118-128.
- Whiting, J.W.M., & Child, I. L. (1953). *Child training and personality: A cross-cultural study*. New Haven, CT: Yale University Press.
- Whiting, B, B. & Whiting, J. W. M. (1975). *Children of six cultures: A psycho-cultural analysis*. Cambridge: Harvard University Press.

Zou, X., Tam, K-P., Morris, M. W., Lee, S-L., Lau, Y-M., Chiu, C-Y. (2008). Culture as *common* sense: Perceived consensus vs. personal beliefs as mechanisms of cultural influence. *Journal of Personality and Social Psychology*, 97, 579-597.

Acknowledgment

This research is based upon work supported in part by the U. S. Army Research Laboratory and the U. S. Army Research Office under grant number W911NF-08-1-0144.