11-2005

The Measurement of Implicit Motives in Three Cultures: Power and Affiliation in Cameroon, Costa Rica, and Germany

Jan Hofer
University of Osnabrück

Athanasios Chasiotis
University of Osnabrück

Wolfgang Friedlmeier
Grand Valley State University, friedlmw@gvsu.edu

Holger Busch
University of Osnabrück

Domingo Campos
University of Costa Rica

Follow this and additional works at: https://scholarworks.gvsu.edu/psy_articles

Part of the Psychology Commons

THE MEASUREMENT OF IMPLICIT
MOTIVES IN THREE CULTURES
Power and Affiliation in Cameroon,
Costa Rica, and Germany

JAN HOFER
ATHANASIOS CHASIOTIS
University of Osnabrück

WOLFGANG FRIEDLMEIER
Grand Valley State University

HOLGER BUSCH
University of Osnabrück

DOMINGO CAMPOS
University of Costa Rica

This article examines methodological issues related to the measurement of implicit motives in culturally divergent samples. Implicit motives are seen as basic needs shared by all human beings. However, cross-cultural comparisons are very restricted because many cross-cultural studies on implicit motives with non-Western cultures developed and discussed culture-inherent stimuli. The aim of the study here was to search for a culture-independent set of picture stimuli measuring two basic motives (affiliation and power motive) in three different cultures. Two pretests and one main study were carried out in Cameroon, Costa Rica, and Germany with student and nonstudent samples, respectively, and an extended methodological cross-cultural analysis was conducted. Construct bias, method bias, and item bias that threaten the cross-cultural comparability of findings were addressed. In analyses, unbiased culture-independent sets of picture stimuli were identified that can be used for cross-cultural comparisons of these two implicit motives.

Keywords: Thematic Apperception Test; equivalence; cross-cultural methodology; test bias; implicit motives; affiliation; power

A method frequently used to measure implicit motives is the Thematic Apperception Test (TAT; Murray, 1943) and its modifications. In research on motivation, this method has been widely used in European American cultures (e.g., McAdams & Vaillant, 1982; McClelland & Pilon, 1983). Only a small number of studies have assessed implicit motives in non-Western cultures by using TAT-type picture-story tests (e.g., Hofer & Chasiotis, 2003, 2004; McClelland & Winter, 1969), most of which applied culturally adapted picture stimuli for data collection (Kornadt & Voight, 1970; Lee, 1953). Consequently, findings on implicit motives from samples in various cultures are barely comparable (Holtzman, 1980). These difficulties could explain why current cross-cultural studies using projective methods are so

scarce (Van de Vijver, 2000). In this article, we report a study designed of two pretests and a main study that aimed to develop a valid set of picture cards measuring the intimacy-affiliation motive and the power motive in three cultures—Cameroon, Costa Rica, and Germany—using student and nonstudent samples.

EQUIVALENCE AND TEST BIAS
IN CROSS-CULTURAL RESEARCH

As validity is not a general characteristic of a test (Cronbach, 1971; Wainer & Braun, 1988) and most tests were developed using data assessed predominately from European American individuals, research on the appropriateness of psychological methods among people from different cultural and ethnic groups is indispensable in cross-cultural research. Therefore, the most important issues in cross-cultural research concern the equivalence of measurements and test bias, respectively (e.g., Allen & Walsh, 2000; Poortinga, 1989; Welkenhuysen-Gybels & Billiet, 2002).

Generally, three levels of equivalence are distinguished: construct or conceptual equivalence, measurement unit equivalence, and scalar equivalence (e.g., Allen & Walsh, 2000; Van de Vijver & Leung, 1997; see also Brislin, 1993). To define each level of equivalence, one has to evaluate the presence or absence of bias. The term *bias* is generally used to describe factors that negatively affect the equivalence of measurements across different (cultural) groups. Van de Vijver and Leung (1997; see also Dana, 1962; Van de Vijver & Poortinga, 1997) describe three major types of bias: construct bias, method bias, and item bias.

Construct bias is present when the construct measured is not identical across cultural groups (i.e., respondents from different cultural groups do not ascribe the same meaning to the construct as a whole, such as conceptions of intelligence). Method bias is related to the form of test administration and is—depending on its source of origin—subdivided into administration bias (e.g., different administration conditions and ambiguous test instructions), instrument bias (e.g., familiarity with test settings and methods of assessment), and sample bias (e.g., sampling differences in test-relevant background characteristics, such as level of education). Finally, item bias is based on characteristics of single items (e.g., content or wording is not equivalent). An item is considered biased when subjects with the same underlying psychological construct (e.g., power motivation) from different (cultural) groups react diversely to a given item (e.g., picture card).

Although construct bias and method bias globally affect the meaningfulness of data from cross-cultural studies, item bias locally influences test scores (Van de Vijver & Leung, 1997). The different kinds of bias, especially item bias, are often studied for objective instruments but almost neglected for projective measurements. Such studies are needed to examine psychometrical adequacy and relevance of projective measurements in cross-cultural research (Van de Vijver, 2000; see also Lindzey, 1961).

One could avoid test bias in cross-cultural research by developing culture-specific tests (e.g., Costantino, Malgady, & Rogler, 1988). However, the theoretical approach of implicit motives points to universal features, and culture-specific instruments cannot detect such common features. Therefore, the aim of this study is to combine universal and culture-specific components of implicit motives into a single theoretical and empirical framework (Hofer & Chasiotis, 2004).

MEASUREMENT OF IMPLICIT MOTIVATION

According to McClelland (1985), implicit motives represent highly generalized preferences derived from emotional experiences during prelingual socialization processes. In contrast to explicit conscious motives such as values and personal goals, implicit motives lack symbolic representation (McClelland, Koestner, & Weinberger, 1989). Implicit motives, which are described as latent behavioral dispositions, are best measured by fantasy-based methods because they clearly express themselves in individuals' fantasies (Biernat, 1989; Weinberger & McClelland, 1990). Many studies have used TAT-type picture-story tests to examine the implicit system of human motivation (e.g., Peterson & Stewart, 1993). Assuming the existence of a limited number of basic human needs, research on motivation concentrates, above all, on an individual's needs for intimacy-affiliation (nAffiliation),[1] power (nPower), and achievement (nAchievement). The intimacy-affiliation motive involves both a concern for warm, close relations with others and a concern for establishing, maintaining, or restoring a positive affective relationship with another person or group of people (Heyns, Veroff, & Atkinson, 1958; McAdams, 1992). The power motive is defined as one's desire to influence or have an effect on the behavior or emotions of other people (Winter, 1973). The implicit achievement motive is defined as an acquired disposition to strive for success in competition with some standard of excellence (McClelland, 1985).[2] It can be assumed that such highly generalized preferences have common features across all cultures.

However, beyond questions of cross-cultural comparability, the TAT has received critical attention for its moderate psychometric conditions in general. Mainly, critics emphasized the poor reliability and, consequentially, the questionable validity of projective measurements (e.g., Campbell & Fiske, 1959; Entwisle, 1972; Klinger, 1966; Klinger & McNelly, 1969; Murstein, 1965a; Weinstein, 1969). Although studies show that situational (e.g., Schultheiss, 2001) and individual characteristics (e.g., French & Lesser, 1964) affect the measurement of implicit motives, underlying processes behind fantasy-based measures seem to be more complex than originally assumed (see Tuerlinckx, De Boeck, & Lens, 2002). However, advocates of TAT-type measurements could refute a number of objections raised by critics (Lundy, 1985; McAdams, 1980; Winter, 1991; Winter & Stewart, 1977). Above all, a large number of studies confirm the validity of thematic apperception measures (Hofer & Chasiotis, 2003; McAdams & Vaillant, 1982; McClelland, 1987; Spangler, 1992; see also Meyer et al., 2001).

DEALING WITH BIAS IN CROSS-CULTURAL
RESEARCH ON IMPLICIT MOTIVES

CONSTRUCT BIAS

There is considerable agreement that motives constitute a universal base of human psyche (Weinberger & McClelland, 1990). However, the conceptual equivalence of definitions of motives across cultures has to be elaborated because significant aspects of the motive genesis are shaped by learning experiences with the satisfaction of needs during early socialization processes that might differ across cultures (Kornadt, Eckensberger, & Emminghaus, 1980; McClelland, 1961; see also Keller & Greenfield, 2000; Markus & Kitayama, 1991). In particular, cross-cultural studies on the achievement motive indicate that the definition of the construct might have to be adapted for studies in non-Western cultures (e.g., De Vos, 1968;

Doi, 1982; Kagan & Knight, 1981; Yu, 1996). In this context, functional equivalence of behavior across cultures has to be examined and to be considered a possible source of construct bias (Allen & Walsh, 2000).

In the literature, various statistical procedures to determine the level of equivalence of cross-cultural findings are discussed, for example, to examine the structure underlying an instrument using confirmatory factor analysis, structural equation models, and cluster analysis (see Espe, 1985; Hagger, Biddle, Chow, Stambulova, & Kavussanu, 2003; Steenkamp & Baumgartner, 1981). Such techniques are not applicable for projective methods and may not even always to be able detect prevailing construct bias (Van de Vijver & Leung, 1997). Valuable information on the construct under investigation (e.g., culture-specific characteristics, behavioral correlates) may be gained by collaboration with local experts, by interviews with cultural informants, and by findings from pretests (e.g., Serpell, 1993). These latter methods were applied in our study.

METHOD BIAS

Method bias can only partly be tested by statistical procedures (e.g., low consistency of responses; see Van de Vijver & Leung, 1997). Therefore, it seems advisable to strictly follow guidelines to control possible sources of method bias when developing a proper design for cross-cultural studies. Referring to the field of research on implicit motives, there are several recommendations to prevent the occurrence of method bias in cross-cultural studies, such as standardized administration, detailed instructions, use of fixed scoring rules, and conduction of stability studies (e.g., Smith, Feld, & Franz, 1992; Van de Vijver, 2000; Veroff, 1992). Furthermore, the assessment of test-relevant background characteristics (subject and context variables; e.g., education, sex, and age) is important to rule out alternative interpretations for cross-cultural and intracultural differences in test scores (Van de Vijver, 2000).

ITEM BIAS

Two main approaches are applicable to detect biased items: the judgmental approach and the statistical approach. Up to now, only a smaller number of studies have tried to identify inappropriate items by judgmental procedures, such as screening of the items by cultural experts (e.g., Van Leest, 1997). The majority of studies examines item bias by employing different statistical methods depending on the measurement level of items, number of (cultural) groups, or sample size (see Van de Vijver & Leung, 1997). Two sorts of item bias can be differentiated: uniform bias and nonuniform bias. Although the former one emerges when the effect of bias is constant across all levels of an underlying trait, the influence of the latter depends on the level of the underlying trait (Welkenhuysen-Gybels & Billiet, 2002). As findings from judgmental and statistical approaches do not sufficiently overlap (Engelhard, Hansche, & Rutledge, 1990; Plake, 1980), we used both strategies in this study.

The three forms of test bias discussed above are interrelated. Comparison of data on implicit motives assessed with identical picture stimuli from Zambian and German participants showed that item bias and method bias impede cross-cultural score comparability (Hofer & Chasiotis, 2004). In addition, it was shown that within the Zambian sample, the validity of intracultural comparisons on motive strengths was threatened by method bias because responses to a number of picture cards were associated with participants' level of education and language use. The latter finding indicates that not only sociodemographic aspects (e.g., social and educational status) but also the enculturation status of participants

(i.e., the extent of a person's active and selective integration into his cultural environment; Huber, 1989; Shimahara, 1970) has to be considered a possible source of bias in intracultural comparisons of test scores. This might be of particular importance in studies with members of open and complex societies but also with members of societies that are exposed to rapid social and cultural changes. Depending on the enculturation status of the participants (e.g., individual variations in regard to adoption of cultural norms and values), an instrument might differ in adequacy (Van de Vijver, 2000; see also Okazaki & Sue, 1995).

THE PRESENT STUDY

The aim of this study was to realize an integrated examination of construct, method, and item bias (see Van de Vijver & Poortinga, 1997) to test the validity of TAT measures for cross-cultural comparisons in three cultural groups. According to the design of our study (two pretests and a main study), the method section is subdivided into three major parts. In the presentation of pretests, we will focus on problems related to construct bias and item bias. The analyses of data from the main study will deal, above all, with the detection of biased picture cards (item bias). Detection of method bias will also be addressed.

SELECTION OF CULTURES

The sampling of the two non-Western cultures was based on the consideration that the cultures should differ from Germany in relevant socioeconomic (e.g., Human Development Index [HDI]) and psychological background variables (e.g., general cultural values or norms) to maximize the possibility of cultural differences. Referring to socioeconomic characteristics, values of the HDI (see United Nations Development Programme, 2003) that are calculated on the basis of data on life expectancy, education, and gross domestic product (GDP) per person indicate that the three nations differ from each other in these three essential elements of human life (see Nohlen & Nuscheler, 1993). Whereas Germany (HDI rank 18) and Costa Rica (HDI rank 42) are both listed among the countries with high human development, Cameroon (HDI rank 142) is characterized as showing low human development.

With regard to differences in psychological background variables, Hofstede (2000) reported higher individualism scores (from 1 to 100) for Germany (67) compared to Costa Rica (15). Cameroon was not included but a sample of the West African region, which may represent a rough approximation for Cameroon, yielded an individualism score of 20. In regard to power distance (scores from 1 to 100), Germany (35) and Costa Rica (35) do not differ, but the West African region showed a much higher power distance (77). Besides, Cameroon and Costa Rica were chosen as study samples because a close collaboration with local cultural experts that seems to be essential in cross-cultural research was at hand. As Cameroon is a multi-ethnic nation, we restricted the sampling of participants to ethnic Nso from the northwest province (Anglophone part of Cameroon; see Yovsi, 2003) to control for cultural differences among African participants.

TECHNIQUES USED FOR DETECTION OF ITEM BIAS

For each story, the absence or presence of motive imagery was coded. According to Tuerlinckx et al. (2002), applying such a dichotomization procedure for thematic apperception measures is far from arbitrary because dichotomized scores correspond to the absence or

presence of motive imagery. The loss of information is small as correlations between nondichotomized and dichotomized scores are sufficiently high (e.g., Entwisle, 1972).

As we were depending on relatively moderate sample sizes (Camilli & Shepard, 1994; Dodeen, 2004), Mantel-Haenszel procedure and the loglinear model analysis were used as techniques to identify item bias. Both techniques are conditional procedures that should be preferred over unconditional procedures to detect item bias (Lord, 1980; Mellenbergh, 1982, 1989). The advantage of conditional methods compared to unconditional methods is that they allow controlling for the subject's overall score on all the items measuring the same latent trait. An unbiased item (or a cue card) means that subjects with the same overall score on average have the same score on an individual item irrespective of the culture to which they pertain. Therefore, a significant effect of culture as a main effect or an interaction effect between culture and overall score implies that the item is biased (Mellenbergh, 1989; Van de Vijver & Leung, 1997).

Both methods proved to yield good results in detection of biased dichotomous items (e.g., Holland & Thayer, 1988; Kok, Mellenbergh, & Van Der Flier, 1985; Potenza & Dornas, 1995; Welkenhuysen-Gybels & Billiet, 2002). Whereas Mantel-Haenszel procedure has high power for identifying uniform bias when comparing data assessed from two (cultural) groups, its model is not appropriate for the detection of nonuniform bias, which is, however, relatively rare (Dorans & Kulick, 1986; Van de Vijver & Leung, 1997). We used it to prescreen picture cards in pretests. Loglinear model analysis allows analyzing data from more than two groups (cultures) at the same time and successfully identifies uniform and nonuniform bias (Welkenhuysen-Gybels & Billiet, 2002). This method was applied for the main study.

PRETESTS IN CAMEROON, COSTA RICA, AND GERMANY—PHASE 1

The first data collections were conducted to gain material for discussions and workshops on motives' definition and indicators with collaboration partners in Cameroon and Costa Rica. In addition, we wanted to get initial information on the stimulus pull and relevance of picture cards for data assessment in divergent cultural samples.

METHOD

PARTICIPANTS

The first pretests in Cameroon and Costa Rica were conducted by our local collaboration partners with small, gender-balanced student samples ($n = 25$ per sample) aged between 16 and 47 years ($M = 22.42$; $SD = 5.55$). In Germany, data of 92 university students aged between 17 and 44 years ($M = 24.13$; $SD = 6.11$) were collected.

PROCEDURE

A picture-story test was administered in a group setting (seminar rooms) in Cameroon, Germany, and Costa Rica. A local assistant informed students who voluntarily participated in data collection that they were going to see a number of pictures. Using an instruction

advised by Smith et al. (1992), students were told that they should try to imagine what is going on in the depicted situation and write a story about the people shown in the picture. It was emphasized that there are no right or wrong stories. Participants should write whatever story they like. As a guide for the stories, four questions were spaced out on the answer sheets: What is happening and who are the people? How did the story begin? What are the people thinking about, and how do they feel? How will the story end? After being shown each picture card for 30 seconds, the students were given 5 minutes to write a story on it.

Costa Rican participants composed their stories in Spanish; Cameroonian students wrote in English. In the northwest province of Cameroon, English represents the official language, whereas Lamnso, the native tongue of the Nso, is one of the colloquial languages. Thus, English is the language predominantly used in educational institutions.

INSTRUMENT

The following picture stimuli (taken from McClelland, 1975, and Smith, 1992) were used for assessing data on implicit intimacy-affiliation and power motivation by presenting them in this order: "an architect at a desk," "women in a lab," "ship captain," "a couple by a river," "trapeze artists," and "nightclub scene." Those picture stimuli are widely used in European American research on motivation. It was assumed that it is not necessary to administer culturally adapted stimuli to assess meaningful data on implicit motives in diverging cultural samples (see Hofer & Chasiotis, 2004; Murstein, 1965b). The number of picture cards was fixed in accordance with known guidelines concerning the assessment of implicit motives. It is generally recommended that not more than six picture cards should be administered on a single test session (see Smith et al., 1992).

CODINGS

Trained assistants at the University of Costa Rica translated the Costa Rican stories into English. The stories of the Cameroonian and Costa Rican participants, together with data on identical picture stimuli that were assessed during the preparation of the study from 92 German university students, established a basis to examine possible sources of bias. Their participation in the study was voluntary. Trained research assistants administered the picture-story test in group sessions using the same instruction as in Cameroon and Costa Rica.

Besides examining initial details on the adequacy of picture stimuli in non-Western samples, the main emphasis lay on an evaluation of indicators of motive imagery scored for the intimacy-affiliation motive and power motive according to Winter (1991). Concern for intimacy and affiliation are coded for any expression indicating positive feelings toward others, sadness about the disruption of a relationship, activities of affiliation, or friendly nurturant acts. Concern for power is coded for responses indicating the impact, control, or influence that a person or social institution has on another person, group, or the world at large (e.g., the forceful determination of behavior, the provision of advice, or impressive displays).

The running text scoring system developed by Winter (1991) can be applied to written and spoken material. In the case of written material such as picture stories, the individual sentence constitutes the scoring unit for motive imagery. In principle, any sentence can be coded for motive imagery (e.g., a certain action, wish, or concern); however, several precise scoring rules have to be considered (e.g., recurrence of motive imagery, negations and questions; see Winter, 1991). All stories assessed during first pretests were coded by the first author and two trained German assistants, who all achieved percentage agreements of 85%

or higher in their responses to training material prescored by experts (Winter, 1991). Scoring disagreements were resolved by discussion to determine the participant's final score.

RESULTS

FINDINGS FROM WORKSHOPS IN CAMEROON

Because of lack of experience with methods for measurement of implicit motivation in Cameroon, in our workshops with Cameroonian assistants we predominantly focused on the appropriateness of constructs' operationalizations: for example, motives' imagery coded for the intimacy-affiliation motive and power motive, respectively, as defined by Winter (1991). Motives' coding and suitability of picture cards were also discussed with Cameroonian study participants to get additional information. In sum, the applicability of the stimulus material for data collection in Cameroon was positively evaluated by cultural experts. Also, a first survey of the cue strength of the picture cards did not point to problems in data collection.

Table 1 gives an overview about the cue strength of the six picture cards used in first pre-tests in Cameroon, Costa Rica, and Germany. It is shown that the two motives are elicited to a sufficient degree in the three cultural samples. Even if a number of noticeable differences in the percentages of elicited motive imagery by a given picture card are present between the subsamples (e.g., higher stimulus pull for nPower of the ship captain in the Cameroonian sample), such differences do not a priori suggest test bias because the distribution of motive strengths may be different in the samples.

Furthermore, it was concluded that the four basic forms of intimacy-affiliation imagery (see Winter, 1991) seem to be adequately assessing nAffiliation among ethnic Nso in Cameroon. Referring to nPower, positive agreement was reached for five of the six indicators of the motive (e.g., having impact on others by strong, forceful actions and control, influencing and impressing behavior). However, it was decided to scrutinize the motive indicator giving help, advice, or support that is not explicitly solicited in more detail to examine whether non-Western participants do perceive this Western power imagery subcategory as being distinct from friendly nurturant acts, which is an imagery subcategory for nAffiliation. In addition, in discussions on the constructs, the question was raised whether an individual's achievement-oriented behavior might be an indicator of nAffiliation when it is centered on benefits of close family and group members.

Evaluations of Cameroonian students. To clarify those issues, we asked Cameroonian high school students to assign possibly underlying motives and goals to fictive persons whose behavior was briefly characterized. Two short descriptions were given to the subjects, one about persons who are involved in advising and supporting other people and the other about someone who is trying to achieve something for the benefit of other people. In three high schools, 97 male and 106 female subjects who were aged between 17 and 30 years ($M =$ 19.06; $SD = 1.52$) participated. As Emmons and McAdams (1991) reported evidence that implicit motives and personal strivings tap common variance, students' responses were categorized using the manual for the coding of personal strivings (Emmons, 1999). We were focusing on the striving categories achievement, combined affiliation-intimacy, and power. The first 35 questionnaires were rated by two scorers. As an interrater reliability of at least .87 for the single categories was reached, the remaining data were coded separately.

TABLE 1
Cue Strength of Picture Cards Used in Pretests Phase I (*n* = 142)

	nAffiliation			nPower		
	Cameroon	Costa Rica	Germany	Cameroon	Costa Rica	Germany
Architect at a desk	56.0	37.5	31.9	52.0	29.2	27.5
Women in a lab	12.0	8.3	15.6	48.0	62.5	57.8
Ship captain	24.0	32.0	20.9	96.0	60.0	76.9
Couple by a river	80.0	80.0	87.6	24.0	28.0	29.2
Trapeze artists	17.4	29.2	38.0	43.5	45.8	54.3
Nightclub	76.2	58.3	69.2	47.6	33.3	60.4

NOTE: Figures represent percentage of subjects answering with at least one indication of a given motive.

The responses of the students to the advice and support description clearly indicated that Cameroonians differentiate between motives that people pursue by such activities. Within the 493 reasons[3] stated by the participants, power strivings (*n* = 200; 40.6%) were considered to be the main motive for such behavior (e.g., to achieve a name in society, be popular, make others follow the right path, etc.). Affiliation-intimacy strivings represent the second main category (*n* = 121; 24.5%). Responses were coded for affiliation-intimacy when a warm concern for the well-being of others and smooth interpersonal relationships were clearly present (e.g., being on good terms with friends, relatives, and neighbors; loving and caring people; to love and to be loved, etc.). Finally, students also named achievement-related strivings (*n* = 48; 9.7%; e.g., to develop the standard of living in society). Referring to the benefit of others description, the same striving categories were used to group the 507 responses of the students. Approximately one fourth of the answers were coded for power (*n* = 141; 27.8%). In addition, 19.9% of the answers were scored for achievement (*n* = 101), and 18.7% of the responses were assigned to the category affiliation-intimacy (*n* = 95).[4] Whereas power codings were, above all, based on status-oriented strivings, strivings for achievement were mostly associated with general development of the community and/or society. Most interesting, an individual's achievement-related behavior seems also to be motivated by affiliation-intimacy-oriented strivings (e.g., concern for others, wish to be part of group, love of other people). This is in line with findings that the distinction between achievement and affiliation motives may be less clear among individuals with an interdependent self-construal (Church & Katigbak, 1992; Maehr & Nicholls, 1980; Yu, 1996).

From a subgroup of the students (*n* = 112), data on uniqueness, which is a subscale of the three-component individualism scale (Realo, Koido, Ceulemans, & Allik, 2002), were assessed. Uniqueness represents a person's awareness of being different from others (see Singelis, Triandis, Bhawuk, & Gelfand, 1995). Realo et al. (2002) reported that uniqueness significantly correlates positively with openness to change and negatively with conservation, both representing higher-order values proposed by Schwartz (e.g., Schwartz et al., 2001; see presentation of main study). To examine whether participants' ratings of the description showed an association with their level of uniqueness (Cronbach's α = .61), regression analyses (method: simultaneous with *z*-transformed variables) were used with level of uniqueness as dependent variable and evaluations of descriptions as being a striving for power, affiliation, and achievement as predictor variables. None of the predictors was significantly related to participants' level of uniqueness.

To conclude, the findings indicate that the use of Winter's (1991) scoring system for implicit motives is warrantable for data collection among Cameroonian Nso. Participants differentiated between power-oriented advice and affiliation-oriented support. Also, achievement-oriented behavior may be an indicator of affiliation-intimacy motivation when its main emphasis is placed on interpersonal relationships and group coherence. However, we assume that two of Winter's affiliation motive imageries (expression of positive, friendly, or intimate feelings toward other people; friendly nurturant acts) are able to cover such motive indicators in the course of a culture-sensitive process of coding.

FINDINGS FROM WORKSHOPS IN COSTA RICA

The workshops in Costa Rica took place after the first field stay in Cameroon. Along with both the German data and the findings from Cameroon, results from the first pretests in Costa Rica were examined with collaboration partners and study assistants at the University of Costa Rica. In brief, it was agreed on the applicability of picture stimuli for data assessment among Costa Rican participants (see Table 1). Furthermore, motives' definitions and motives' imageries were examined and regarded as being adequate for measurement of implicit motives in Costa Rica. Referring to Cameroonian findings, it was also decided to adhere to affiliation-oriented achievement imagery. Because Costa Rican cooperation partners were familiar with thematic apperception methods and had used them in collecting data on motivation (e.g., Campos, 1991), no additional inquiries on motive indicators were organized.

To avoid problems related to translation of transcripts, it was planned that local assistants will score the stories in future data collections. Thus, five Costa Rican assistants were trained in Winter's scoring guidelines of implicit motives until they had achieved percentage agreements of 85% or better with prescored training materials taken from the scoring manual (Winter, 1991).[5]

PRETESTS IN CAMEROON, COSTA RICA, AND GERMANY—PHASE 2

The main goal of the second phase of pretests was an analysis of item bias to put together a set of culturally unbiased pictures for the assessment of needs for affiliation and power, respectively, in the main study. Picture cards were applied to larger student samples. It was decided to use two additional cards for the assessment of nAffiliation and nPower to have a wider scope of choice if bias analyses based on culture point to shortcomings of some picture stimuli within the set.

METHOD

PARTICIPANTS

In Cameroon, 98 subjects from a teachers training college—namely, 52 female and 46 male students aged between 19 and 45 years ($M = 22.50$; $SD = 2.98$)—took part in this study.

The Costa Rican study sample consisted of 99 (75 female and 24 male) university students (psychology and economics, respectively) who were between 17 and 50 years old ($M = 23.23$; $SD = 5.49$). Finally, in Germany 73 (49 female and 24 male) participants who were studying several subjects (e.g., geography, education, and psychology) were recruited at three universities. The German students were aged between 19 and 65 years ($M = 24.97$; $SD = 8.00$). All students voluntarily participated in the study.

PROCEDURE

As in first pretests, the picture-story test was administered by local assistants in group settings (seminar rooms), using the identical instruction and time frame. To examine order effects on the measurement of social motives, the picture stimuli were administered in two different orders.

INSTRUMENT

As mentioned above, we added two cards that were used by Veroff and colleagues (Veroff, Feld, & Crockett, 1966): "man with a cigarette behind woman" and "four men seated at a table" (both cards are reproduced in Smith, 1992).

CODINGS

The picture stories were coded following the guidelines set forth in first pretests by trained German (German and Cameroonian data) and Costa Rican assistants (Costa Rican data). At first, stories of 20 Costa Rican participants were translated by a bilingual assistant at the University of Osnabrück and coded by German and Costa Rican assistants to determine scoring disagreements. The interrater reliabilities (percentage agreement) were .85 for both nAffiliation and for nPower.[6] Costa Rican assistants received substantial feedback on scoring disagreements before they continued the coding of local participants' stories.

DATA ANALYSIS

Mantel-Haenszel procedure was applied to test the item bias. This procedure tests whether the percentages of motive imagery for a picture card are identical in two groups (cultures) for all score levels (Van de Vijver, 1994; Van de Vijver & Leung, 1997). Scores were dichotomized as follows: If no imagery for a given motive was present, a new score of 0 was assigned, and if at least one imagery was coded, a new score of 1 was assigned. Three evenly sized score levels were differentiated: low, medium, and high strength of motive (see Van de Vijver & Leung, 1997). However, for each single bias analysis, the motive strength was recalculated because the item scrutinized for bias should not contribute to the ability index. Analyses were carried out for each motive separately based on three cultural comparisons (Germany-Costa Rica, Germany-Cameroon, and Cameroon-Costa Rica). A 2 (culture) × 2 (score group) matrix was composed for each item (picture). A significant chi-square value indicates that a picture cue is biased and no valid (cross-cultural) group comparisons can be performed. As recommended (e.g., Lord, 1980; Van Der Flier, Mellenbergh, Adèr, & Wijn, 1984), we used an iterative strategy in bias analyses: After the identification of the most

significantly biased item, new score levels and score groups were determined. The procedure was repeated until the set was free of items that are significantly biased (see Van de Vijver & Leung, 1997; Welkenhuysen-Gybels & Billiet, 2002).

RESULTS

For further analyses, only data of 265 study subjects were used because 5 participants did not write stories to the complete picture set. All remaining participants produced enough written material for scoring (averaging 30 words per story; see Smith et al., 1992). The total number of words ranged from 262 to 1,091 words ($M = 598.8$; $SD = 142.7$) in the total study sample ($N = 265$). Story scores for motive imageries were significantly correlated with story length (affiliation: $r = .39$; power: $r = .31$; ps < .01). Dividing raw scores by the length of the stories does not always correct for length because the correlation of length with the resulting quotient scores is not necessarily zero. This given, the effects of correlations between motive scores and story length were removed, employing a regression method (Winter, 1992). New scores were calculated as follows: actual motive score minus the protocol length multiplied by the constant (i.e., correlation coefficient of length with motive score multiplied by the ratio of standard deviations of motive scores and protocol length; see Winter, 1992). The correction was computed for each story individually (Smith et al., 1992). Item-corrected scores were added to new sum scores. No picture order effects on stimulus pull and motive strength could be verified.

INITIAL ANALYSIS OF ITEM BIAS

Referring to nAffiliation, in the first step of item analyses with all picture cards scrutinized for bias, it was indicated that the "architect at a desk" was the most significantly biased item. Costa Rican participants reported a motive imagery significantly more often across all score levels than German ($\chi^2 = 14.40$; $p < .01$) and Cameroonian participants ($\chi^2 = 12.78$; $p < .01$) (see Table 2). In the second step, "man with cigarette" was removed. Cameroonian participants showed higher motive imagery than German ($\chi^2 = 17.11$; $p < .01$) and Costa Rican participants ($\chi^2 = 6.01$; $p < .05$). Finally, "women in a lab" was removed because Costa Ricans scored significantly more often than Cameroonian students ($\chi^2 = 4.50$; $p < .05$). The remaining five picture cards proved to be free of bias in further analyses.

Analyses on nPower revealed that initially "women in a lab" should be removed. Cameroonian showed a higher amount of motive imagery across all score levels than Costa Rican ($\chi^2 = 17.56$; $p < .01$) and German students ($\chi^2 = 10.33$; $p < .01$). In the next step, "man with cigarette" was excluded. It evoked the power motive significantly more often for German compared to Cameroonian ($\chi^2 = 8.36$; $p < .01$) and Costa Rican participants ($\chi^2 = 5.91$; $p < .05$). The "ship captain" was removed in a next step: Cameroonian had significantly higher values than Cost Rican students ($\chi^2 = 8.57$; $p < .01$). In further analyses on nPower, no more indication of item bias was found.

As we were aiming to identify a uniform set of picture cards applicable for the measurement of the two given motives, the "ship captain" and the "architect at a desk," respectively, were removed for further analyses on both nAffiliation and nPower. In subsequent analyses, no indicator of item bias could be verified. Thus, the unbiased set for further testing consists of four cards (see Table 2)—namely, a couple by a river, trapeze artists, nightclub scene, and four men seated at a table.[7]

TABLE 2
Item Bias Analysis of Cue Strength of Picture
Cards Used in Pretests Phase 2 ($N = 265$)

	nAffiliation			*nPower*		
	Cameroon	*Costa Rica*	*Germany*	*Cameroon*	*Costa Rica*	*Germany*
Unbiased picture cards						
A couple by a river	70.4	86.9	83.1	14.3	12.1	14.1
Trapeze artists	12.2	31.3	18.3	38.8	44.4	57.7
Nightclub	44.9	72.7	64.8	37.8	43.4	53.5
Men at a table	28.9	32.3	41.4	32.0	33.3	44.3
Biased picture cards						
Architect at a desk	30.6 <	58.6 >	26.8	18.4	20.2	21.1
	30.6	=	26.8			
Women in a lab	4.1 <	12.1 >	8.5	30.6 <	59.6 =	60.6
	4.1	=	8.5	30.6	<	60.6
Man with cigarette	50.0 >	35.7 =	21.4	54.1 =	61.2 <	82.9
	50.0	>	21.4	54.1	<	82.9
Ship captain	16.3	23.2	12.7	78.6 >	59.6 =	74.6
				78.6	=	74.6

NOTE: Figures represent percentage of subjects answering with at least one indication of a given motive. For biased items, the results of the three comparisons based on the Mantel-Haenszel procedure are reported. The < and > symbols indicate significant differences.

MAIN STUDY

The main study aimed to test the appropriateness of the selected picture set based on the pretests. In addition, participants' value orientations were assessed to identify enculturation differences between both cultures and individuals. It was intended that each cultural sample be composed of participants that are characterized by a wide range of important background characteristics such as gender, age, education, and socioeconomic status. The three cultural samples should be balanced concerning these sample characteristics. Therefore, three non-student samples from Cameroon, Costa Rica, and Germany were studied.

METHOD

PARTICIPANTS

Trained local assistants collected data from 370 participants in Cameroon ($n = 126$), Costa Rica ($n = 120$), and Germany ($n = 124$). Each sample was balanced with respect to rural/urban distribution of participants. Furthermore, within each regional context, samples were gender balanced. Finally, the distribution of female and male participants, respectively, across age groups (18 to 25 years, 26 to 45 years, 46 years and older) was kept equal in rural and urban areas. In Cameroon, participants' ages ranged from 20 to 65 years ($M = 35.88$; $SD = 13.91$). Costa Rican subjects were between 18 and 74 years old ($M = 36.39$; $SD = 15.09$), and German participants' ages ranged from 18 to 75 years ($M = 36.94$; $SD = 13.80$). Accordingly, the three cultural samples did not significantly differ in age. It was also intended to recruit

participants with differing educational levels and occupational activities in each culture to diversify the socioeconomic status of the subjects. Referring to level of education, 45.4% of the total sample ($N = 168$: 72 Cameroonians, 57.1%; 39 Costa Ricans, 32.5%; 57 Germans, 46.%) were categorized as having a low level of education (less than secondary school education), and 54.6% of the study participants ($n = 202$: 54 Cameroonians, 42.9%; 81 Costa Ricans, 67.5%; 67 Germans, 54%) were assigned to the category high level of education (secondary school education and more).

PROCEDURE

Cameroonian and Costa Rican participants were visited at their homes for data collection. In Germany, measurements were administered to the majority of participants on the premises of the university. German and Cameroonian people received a monetary compensation for their participation. In contrast, compensation was not given to Costa Rican participants. We were informed by our collaboration partners that Costa Rican people do not expect it and might feel irritated when compensated for voluntary participation.

In general, participants received the same instruction as student samples in pretests. As we were collecting data on implicit motives from nonstudent samples, the instruction was, however, slightly modified for all cultural groups (e.g., more detailed, giving of an example; see DeCharms, 1992). Also, the conditions of data collection were modified as a consequence of the recruitment of nonstudent participants. In contrast to pretests, the test methods were administered to participants individually by local research assistants. Furthermore, the assessment of picture-story tests could not be carried out in the standardized way in Cameroon. Twenty-six subjects responded verbally and were recorded because of insufficient penmanship. Those stories were transcribed and coded.

Having finished the picture-story test, participants in the three cultures completed the Schwartz Value Survey (SVS; Schwartz, 1992), which was given to them in (official) first-language versions (see e.g., Bilsky & Peters, 1999; Schwartz, 1992; Schwartz & Bilsky, 1990).

INSTRUMENTS

The four cross-culturally valid pictures based on the results of the pretest—a couple by a river, trapeze artists, a nightclub scene, and four men seated at a table—were used. As the pretest results for data from Cameroon and Costa Rica indicated that nightclub scene shows a trend to be biased with respect to nAffiliation ($\chi^2 = 3.30$; $p < .10$), a further picture stimulus that depicts a meaningful situation across cultures was included into the set: "man and children seated at a table" (see Smith, 1992).

The SVS is widely used in cross-cultural research on values (guiding principles in life; e.g., Schwartz & Sagiv, 1995). The items of the questionnaire represent 10 value types (e.g., achievement, benevolence, conformity, and self-direction) that on a more abstract level reflect two bipolar dimensions (higher order value types; Schwartz, 1994a): openness to change (self-direction and stimulation) versus conservation (conformity, security, and tradition) and self-enhancement (achievement and power) versus self-transcendence (benevolence and universalism). According to Triandis (1995, 1996), openness to change overlaps with individualism and conservation with collectivism. The SVS was used to test the assumption that participants from Cameroon, Costa Rica, and Germany differ in psychological characteristics such as value orientations. To determine participants' values on higher-

order dimensions, only items that were found to have similar meanings across cultures were included (Schwartz, 1994a).[8] Most of the higher-order value types show satisfactory to high internal consistencies across the total sample but also within each cultural sample (Cronbach's alphas > .70). Only among Cameroonian participants, openness to change (.58) shows an alpha that falls just below an acceptable value of .6 (Nunnally, 1978).

CODINGS

Like in pretests, stories were coded by trained German (German and Cameroonian data) and Costa Rican assistants (Costa Rican data). Again, 30 picture-story tests from Costa Rican participants that were translated by a bilingual assistant in Germany were coded by Costa Rican and German assistants to examine the reliability of raters. The interrater reliability was slightly better than in pretests: .86 for nPower and .91 for nAffiliation. Costa Rican assistants received feedback on scoring disagreements before they scored the remaining data.

DATA ANALYSIS

One-way ANOVAs were used to test cultural differences in cultural values. Loglinear models were used for item bias detection. Loglinear modeling allows testing the fit of different, nested models (one-way or two-way effects) to the observed frequencies in the cross-tabulation of categoric variables. Among hierarchically related models, the most parsimonious model that fits the data is preferable. Here, participants' dichotomized motive scores (0 = no coding; 1 = coding) for nAffiliation and nPower, respectively, were used as dependent variables.[9] Culture and score level were the two factors. Three score levels were differentiated: low, medium, and high.

A saturated model that would always exactly match the observed frequencies with no degrees of freedom remaining includes all possible one-way and two-way effects—that is, score level (low-medium-high), culture, and interaction of culture and score level. Therefore, a more parsimonious model that adequately fits the data must be identified. An item shows no nonuniform bias when a model that includes score level and culture does fit the data. Furthermore, a fitting model that includes only score level is an additional indicator of absent uniform bias (Van Der Flier et al., 1984). Likelihood ratio chi-square was used as a test statistic for evaluating the adequate model. Similar to the analysis of Pretest 2, an iterative strategy was applied. Further loglinear models were analyzed by taking intracultural characteristics (age, gender, level of education, and value orientation) into account.

RESULTS

CROSS-CULTURAL COMPARISONS OF GUIDING PRINCIPLES IN LIFE

One-way ANOVAs for "culture," accompanied by a Scheffé post hoc test, yielded clear differences in value orientations between cultural groups as expected: Cameroonian subjects reported a lower importance of openness to change ($F_{2, 361} = 60.18$; $p < .01$; $M = 2.82$; $SD = 1.10$) than did participants from Costa Rica ($M = 3.98$; $SD = 1.11$; $p < .01$) and Germany ($M = 4.21$; $SD = .96$; $p < .01$). Referring to conservation, participants from Germany markedly differed from Cameroonian and Costa Rican subjects ($F_{2, 361} = 116.52$; $p < .01$). German

subjects placed less emphasis on conservation ($M = 3.20$; $SD = .84$) than did Costa Rican ($M = 4.53$; $SD = .99$; $p < .01$) and Cameroonian subjects ($M = 4.78$; $SD = .75$; $p < .01$). Furthermore, the comparison between Cameroonian and Costa Rican subjects also tended to be significant ($p < .08$). German subjects ($F_{2, 364} = 15.36$; $p < .01$; $M = 2.67$; $SD = .93$) rated self-enhancement as being significantly less important as a guiding principle in their life than did Cameroonian ($M = 3.34$; $SD = 1.16$; $p < .01$) and Costa Rican subjects ($M = 3.28$; $SD = 1.02$; $p < .01$). Finally, Costa Rican participants ($F_{2, 361} = 10.54$; $p < .01$; $M = 5.13$; $SD = .62$) rated the importance of the higher order value type self-transcendence significantly higher than participants from Cameroon ($M = 4.75$; $SD = .81$; $p < .01$) and Germany ($M = 4.72$; $SD = .82$; $p < .01$), respectively.

SCREENING OF MOTIVE MEASUREMENTS

First, it was checked whether all participants produced enough material to permit a meaningful scoring of the stories. The data of 8 German and 2 Costa Rican participants were excluded from further analyses because their averaged story length clearly fell below the critical value (see Smith et al., 1992). Two participants from Cameroon responded to picture cards using pidgin English. To control for language as a possible source of bias with the Cameroonian subsample (see Ervin, 1964; Hofer & Chasiotis, 2004), data from those two participants were not considered for further analyses. Among the remaining study sample ($N = 358$), the number of words for the five stories ranged from 149 to 1,023 ($M = 316.31$; $SD = 129.96$). The stories were coded by German and Costa Rican assistants following the guidelines set forth by Winter (1991). In addition, attention was paid to affiliation-oriented achievement images. However, such behavior was rarely present in the stories of the study participants; only two Cameroonian stories were coded for such motive imagery (e.g., intention to bring people closer to each other by developing environmental conditions).

To give a first impression on data on implicit motives assessed from nonstudent samples, Table 3 presents cue strengths of the five picture cards in Cameroon, Costa Rica, and Germany ($N = 358$). In general, the stimulus pull of the picture cards seems to be similar to the ones found in pretests (see Table 2). Like in pretests, motive scores were corrected using a regression method (Winter, 1992) because of significant correlations between protocol length and number of motive images ($r = .28$ for nAffiliation and $r = .35$ for nPower; $p < .01$).

ANALYSES OF BIAS

Motives' measurement in Cameroon. In a first step, the results of the stories of 26 subjects who answered verbally were compared with the results of the other Cameroonian subjects to test whether the different method biased test scores. In comparisons of total motive strength of nAffiliation and nPower, no differences between both Cameroonian subgroups could be verified. Also, in bias analyses no indicator of differential responses of participants from both subsamples to picture cards was found. No biased picture cards were identified with respect to affiliation and power motivation. Thus, Cameroonian participants that verbally responded to picture stimuli were not excluded from further analyses because no evidence for method bias seems to be present.

Item bias analysis based on culture and score level. Referring to nAffiliation, Table 4 gives an overview of the first run of bias analyses. Three cards show uniform bias: a couple

TABLE 3
Cue Strength of Picture Cards Used in Main Data Collection ($N = 358$)

	nAffiliation			nPower		
	Cameroon	Costa Rica	Germany	Cameroon	Costa Rica	Germany
A couple by a river	63.7	92.4	87.1	13.7	9.3	11.2
Trapeze artists	29.0	32.2	23.3	34.7	36.4	32.8
Nightclub	51.6	75.4	69.0	24.2	28.8	40.5
Men at a table	29.8	44.1	37.1	34.7	26.3	34.5
Man and children	24.2	46.6	31.0	63.7	61.9	87.9

NOTE: Figures represent percentage of subjects answering with at least one indication of a given motive.

by a river ($\chi^2 = 29.09$; $p < .01$), nightclub scene ($\chi^2 = 16.86$; $p < .05$), and man with children ($\chi^2 = 14.08$; $p < .05$). In addition, test statistic indicates nonuniform bias for men at a table ($\chi^2 = 10.45$; $p < .05$). In the next step, the picture set was reduced by the most biased item (highest chi-square; a couple by a river). Ability indexes and score level groups were recalculated and analyses were repeated. None of the remaining four picture cards proved to be biased (uniform and nonuniform) in analyses that followed.

The same procedure was followed for implicit power motive (see Table 5). In analyses, none of the cards was identified for nonuniform bias. Also, only man with children showed evidence for uniform bias ($\chi^2 = 31.37$; $p < .01$). The elimination of this item resulted in a set of four picture cards that showed no evidence of either uniform bias or nonuniform bias. The motive strengths measured by the unbiased four-picture sets and the originally five-picture set correlate significantly with each other—that is, .91 ($p < .01$) for nAffiliation and .86 ($p < .01$) for nPower.[10]

Additional analyses on context variables. To test the quality of the emerged four-picture sets, further analyses were conducted following Okazaki and Sue (1995), who noted that ethnicity as a demographic variable might not be the variable of interest in cross-cultural research. Rather, variables that cover psychological constructs associated with culture may be relevant (e.g., cultural orientations). Such a substitution of culture with psychological variables, which are related to both dimensions of individualism and collectivism, may also provide evidence for the occurrence of method bias associated with the enculturation status of the participants in data assessment (Van de Vijver, 2000).

In the following, it was tested by the use of loglinear models whether the picture sets proved to be unbiased concerning participants' reported importance of higher order values of conservation and openness to change, respectively (low vs. high importance). In additional analyses, the influence of participants' age (three age groups), gender, and level of education (low vs. high educational level) were examined. In analyses, a given picture card's dichotomized data for nAffiliation and nPower, respectively, was entered as a response variable. Participants' score level of a given motive (low-medium-high),[11] one of the context variables (e.g., openness to change), and the interaction of both formed the set of independent variables. With respect to nPower, none of the picture stimuli proved to be biased in analyses on participants' gender, age, and the reported level of both conservation and openness to change. One of the analyses on nPower showed a significant main effect of participants' level of education (uniform bias). Across the three score levels, highly educated participants were coded significantly more often for nPower in responses on nightclub scene ($\chi^2 = 12.07$; $p <$

TABLE 4
nAffiliation: Initial Item Bias Analysis (Loglinear Models) of
Cue Strength of Picture Cards in the Main Study ($N = 358$)

Motive		*nAffiliation*				
Picture Card		*Strength of Motive (%)*[a]			*Item bias*	
Culture	*n1, n2, n3*[b]	*Low*	*Medium*	*High*	*Uniform*	*Nonuniform*
A couple by a river					$\chi^2(df = 6)$ (p)	$\chi^2(df = 4)$ (p)
Cameroon	62, 31, 31	53.2	71.0	77.4		
Costa Rica	21, 46, 51	95.2	89.1	94.1	29.09**	2.85
Germany	36, 43, 37	80.5	90.7	89.2		
Trapeze artists						
Cameroon	67, 35, 22	23.9	22.8	54.4		
Costa Rica	18, 48, 52	22.7	27.1	40.4	7.46	3.84
Germany	34, 37, 45	8.8	27.0	31.1		
Nightclub						
Cameroon	59, 39, 26	45.7	48.7	69.2		
Costa Rica	22, 39, 57	81.8	74.3	73.7	16.86*	5.83
Germany	38, 42, 36	55.2	71.4	80.5		
Men at a table						
Cameroon	68, 29, 27	20.6	20.7	62.9		
Costa Rica	14, 55, 49	35.7	41.8	49.0	11.70	10.45*
Germany	37, 36, 43	27.0	47.2	37.2		
Man with children						
Cameroon	68, 29, 27	23.5	17.2	33.3		
Costa Rica	16, 56, 46	18.7	44.6	58.7	14.08*	4.81
Germany	35, 35, 46	25.7	31.4	34.8		

a. Percentage of subjects answering with at least one indication of a given motive.
b. Number of participants with low, medium, and high motive strength, respectively, in the three subsamples.
*$p < .05$. **$p < .01$.

.01). No indication of bias was found for the remaining three picture cards. Referring to nAffiliation, no evidence test bias was found in analyses.

DISCUSSION

The main goal of the study was to illustrate how to deal with the issue of equivalence and comparability in the context of cross-cultural research on implicit motives. Subjects in Cameroon, Costa Rica, and Germany participated in this study. In particular, it was examined whether a culture-independent, cross-culturally valid set of picture stimuli can be found in three different cultures that measures two basic implicit motives (affiliation and power motive). Therefore, an integrated examination of construct, method, and item bias was conducted. Such extensive methodological perspective is recommended by Van de Vijver and Poortinga (1997) to improve the validity of cross-cultural comparisons.

DIFFERENCES IN VALUE ORIENTATION

The assumptions about cultural differences in value orientations (e.g., Hofstede, 2000; Singelis et al., 1995; Triandis, 1996) were empirically confirmed. Cameroonian participants

TABLE 5
nPower: Initial Item Bias Analysis (Loglinear Models) of
Cue Strength of Picture Cards in the Main Study (*N* = 358)

Motive		*nPower*				
Picture Card		*Strength of Motive (%)*[a]			*Item bias*	
Culture	*n1, n2, n3*[b]	*Low*	*Medium*	*High*	*Uniform*	*Nonuniform*
A couple by a river					$\chi^2(df=6)\,(p)$	$\chi^2(df=4)\,(p)$
Cameroon	50, 41, 33	10.0	9.7	24.2		
Costa Rica	42, 43, 33	2.4	9.3	18.2	6.09	3.95
Germany	27, 36, 53	0	8.3	18.9		
Trapeze artists						
Cameroon	55, 38, 31	34.5	28.9	41.9		
Costa Rica	43, 39, 36	44.2	28.2	36.1	1.66	1.38
Germany	21, 43, 52	33.3	30.2	34.6		
Nightclub						
Cameroon	48, 40, 36	22.9	12.5	38.9		
Costa Rica	42, 44, 32	26.2	27.3	34.4	8.34	2.42
Germany	29, 36, 51	37.9	30.5	49.0		
Men at a table						
Cameroon	55, 37, 32	29.1	32.4	46.9		
Costa Rica	42, 40, 36	30.9	20.0	27.8	4.58	2.14
Germany	22, 43, 51	31.8	27.9	41.2		
Man with children						
Cameroon	47, 41, 36	61.7	65.8	63.9		
Costa Rica	38, 48, 32	60.5	62.5	62.5	31.37**	5.18
Germany	34, 31, 51	97.0	80.6	86.3		

a. Number of participants with low, medium, and high motive strength, respectively, in the three subsamples.
b. Percentage of subjects answering with at least one indication of a given motive.
*$p < .05$. **$p < .01$.

reported both the lowest level of openness to change and the highest level of conservation among the three cultural samples. In addition, differences in self-transcendence occurred between the three subsamples in which Costa Rican participants placed the highest importance on this higher order value type. At first view, the finding that German participants reported the lowest importance of self-enhancement seems to contradict expected differences between so-called individualistic and collectivistic cultures. However, Schwartz (1992, 1994b; see also Oishi, Schimmack, Diener, & Suh, 1998) could show that even if power- and achievement-related values serve individual interests at the individual level, in collectivistic cultures they are often rated as more important than in individualistic cultures.

To summarize, findings on value orientations indicate that both non-Western samples clearly differ from the Western sample in their reported importance of guiding principles in life. Also, non-Western samples also differ from each other with respect to value orientations.

CONSIDERATION OF CONSTRUCT BIAS

Issues concerning equivalence of constructs measured in the three cultures were addressed in the first phase of pretests with student samples. Because definition and measurement of a given construct are closely interrelated (Eckensberger, 2002), the

cross-cultural meaningfulness of established motive indicators was addressed. Many studies in Western cultures assessed data on implicit motivation by using picture-story tests including Winter's (1991) manual for scoring motive imagery, and Costa Rican cooperation partners were familiar with both. The discussion with cultural informants showed high congruence in regard to the coding rules. Thus, questions related to construct bias were, above all, examined in the Cameroonian subsample because no information on the appropriateness of thematic apperception measurement was available. Following a suggestion by Serpell (1993; see also Super, 1983), a survey with local informants was conducted to examine the validity of those indicators of implicit motives, which were identified as problematic in discussions with local assistants in Cameroon. Two critical coding rules were discovered— "giving advice and/or support" as part of the affiliation instead of power motive and "achievement for others" as affiliation instead of achievement motive. The qualitative analysis of the answers showed that those motive indicators used by Winter (1991) were adequately assessing information on nAffiliation and nPower in Cameroon. It is noteworthy that Cameroonian participants with varying levels of describing themselves as being unique did not differ from each other in evaluations of fictive behaviors (Realo et al., 2002). This result might indicate that behavioral correlates of given implicit motives are relatively independent of cultural orientations reported by an individual. It is interesting that affiliation-oriented achievement behavior was rarely present in the picture stories of participants. Therefore, it is assumed that Cameroonian participants' evaluations of such strivings reflect prevalent, internalized cultural orientations rather than behavioral correlates of implicit affiliation motivation. Findings from Cameroon were introduced in discussions with Costa Rican assistants and considered in subsequent data assessments and scoring.

CONSIDERATION OF METHOD BIAS

Various precautions were set to prevent the occurrence of method bias in data collection. First of all, a number of recommendations set forth in literature (e.g., Smith et al., 1992; Van de Vijver, 2000) were adhered to such as extensive training of local test administrators, use of fixed scoring rules, and examination of interrater agreements. Furthermore, there was an effort to obviate group differences in familiarity with stimulus material and testing situation—a recurrent problem in cross-cultural studies—by giving participants from all cultural groups a detailed and vivid introduction to picture-story test (see Hofer & Chasiotis, 2004). In addition, Costa Rican data were scored by trained local assistants to avoid problems associated with translation of data. Furthermore, data from Cameroonian participants who did not respond in English were excluded from analyses to control for another source of bias (Hofer & Chasiotis, 2004). A standardization of test administration could only be implemented in pretests because in the main data collection, cultural samples were composed of study participants who clearly differed from each other in their level of education. Thus, in Cameroon a number of participants verbally responded to picture stimuli. Analyses within the Cameroonian subsample, however, did not point to bias associated with differences in test administration. Attempts to prevent method bias were further evident in the process of sample recruitment. It was intended to balance cultural samples with respect to important subject and context variables (e.g., gender, age, and education).

Using a number of those variables, data were further screened on potential bias.[12] Although cultural samples markedly differed in cultural orientations (higher order values; Schwartz, 1992), no indicator of systematic bias was found in analyses based on participants' level of both openness to change and conservation.

Furthermore, bias analyses were conducted based on participants' age, gender, and level of education. None of these analyses but one did point to evidence of bias. However, across the three score levels, higher educated participants were coded more often for nPower in their stories on the nightclub scene than subjects with a low level of education. As method bias usually affects scores at the level of the whole instrument (all picture cards), we do not consider this finding evidence for that kind of bias. Furthermore, separate analyses for each cultural group on the level of education and power motivation did not produce evidence for bias of that picture card. This latter finding leads us to conclude that we should not overestimate the negative impact on the validity of data on nPower in Cameroon, Costa Rica, and Germany.

CONSIDERATION OF ITEM BIAS

Item bias analyses based on culture were conducted with data assessed during both pretests and main study. In a first step, the applicability of picture material was discussed with local experts (judgmental approach), and picture cards were screened for stimulus pull among participants from the three cultural student samples. Even if cultural experts agreed that preselected picture cards would generally be applicable in non-Western cultures, statistical analyses on differential item functioning were conducted in the next step. In analyses on data assessed during the second phase of pretests, 50% of the picture stimuli (four cards) were identified as biased. Together with an extra added card, the remaining four picture cards formed the set of stimuli for the main data collection. The quality of the picture set was further examined using data assessed from samples that were extended in regard to educational level, socioeconomic status, and nonstudent samples to broaden the intracultural variance. It was possible to identify bias-free picture cards that can be used for cross-cultural comparisons of nAffiliation and nPower, respectively. In this context, we have to mention that the stimulus pull for nPower of the developed set is considered to be moderate because the stimulus showing the highest pull across the three cultures was biased.

In sum, we conclude that the identification of fairly unbiased picture sets for the measurement of implicit affiliation and power motivation among study participants from three cultures was successful. Our results indicate that the final picture cards may be selected in a first step as promising stimulus cues for future assessments of implicit needs for affiliation and power among similar cultural samples. But additional statistical screening on test bias is indispensable because shortcomings of the present study must be considered.

LIMITATIONS AND FUTURE PERSPECTIVES

Among these limitations, statistical procedures used for detection of item bias have to be mentioned. Methods that we have used for the detection of item bias all compare performance levels across total score levels. Van de Vijver and Leung (1997), however, noted that different procedures often differ in findings on biased items (see also Rudner, Getson, & Knight, 1980). A further problem is the low stability of item bias statistics. Thus, techniques different from the ones we have used might produce different results (see, e.g., Tuerlinckx et al., 2002). Generally, an important problem in conditional item bias techniques that use a sum score as ability index (with or without the score of the item screened for bias) is the number of biased items. The significance of the sum score as an adequate estimation of the strength of the underlying construct depends on the presence of few biased items (Van de Vijver, 1994). Also, the removal of biased items may invalidate the latent trait's estimator.

This may particularly be a problem if one has to rely on a few items or picture cards. It is advisable to use not more than four to six stimuli assessing data on implicit motives by using picture-story test because additional stories may be less valid than the earlier ones (see Smith et al., 1992).

Alternatively to methods that we implemented for analysis, different techniques (e.g., Swamanithan & Rogers, 1990; Welkenhuysen-Gybels, 2003) or models derived from item response theory (e.g., Borsboom, Mellenbergh, & van Heerden, 2002; Welkenhuysen-Gybels & Billiet, 2002) may be applied. These models allow testing the bias for the whole set of pictures at once instead of separate analysis for each item; for example, the specific objectivity (i.e., the order of the extent of evoking the respective motive) is the same for all three cultural groups (see also Millsap & Everson, 1993; item impact). It is a task for the future to apply such a method to evaluate their advantages and disadvantages. Generally, detection of item bias across more than two (cultural) groups has received little attention (Welkenhuysen-Gybels, 2003). Thus, further (simulation) studies are needed before the pros and cons of different methods are identified fully.

Generally, it is important to mention that identification and subsequent elimination of biased items does not guarantee valid data because there are more threats to the validity of group comparisons in cross-cultural research (Holland & Wainer, 1993; Van de Vijver, 1994). Participants' differing familiarity with both test methods and test situation and the problems associated with the use of different languages are commonly listed among factors that negatively affect comparability of cross-cultural data. In spite of paying attention to both sources of bias, we cannot entirely exclude negative effects of such factors on the quality of data. Concerning analyses on context variables by which we further tested the quality of the bias-free picture sets, the selection of these variables was, of course, limited. So it might be the case that we have missed important variables that would have helped to identify bias inherent in the data. Furthermore, it has to be mentioned that additional data assessments with respect to construct bias were only performed in Cameroon. This procedure is a good example that cross-cultural analysis forces researchers to be more sensitive in regard to methodological issues compared to other researchers. It would also be interesting to do the same test with Costa Rican and German subsamples.

One could argue that picture cards depicting Western individuals may not be useful for cross-cultural research. In our opinion, it is not necessary to use culturally adapted picture cards for cross-cultural research (see Hofer & Chasiotis, 2004). However, verbal stimuli may be a possible alternative as stimulus material to elicit motive imagery in cross-cultural studies.

Whereas we found evidence that a number of unbiased picture cards can be used for cross-cultural comparisons on motive strength, this study also demonstrates that marked cross-cultural differences exist concerning situations that are cross-culturally comparable with respect to motive arousal and realization. Half of the used picture cards in Pretest 2 showed a bias. Even if individuals may have a general desire for affiliation and power, these needs might be accomplished in different situations and in various ways. Thus, more attention should be given to the differentiation of the motive (end) and the possibly culture-bound ways that a motive is realized (means; see Maehr, 1974) because in how far different aspects of implicit needs will develop depends on what kind of opportunities, incentives, and encouragement is provided within a given sociocultural context (see Kornadt et al., 1980). A contextual analysis of motive-related behavior in non-Western cultures seems to be promising to identify situations that are characterized by cues that signal positive or negative incentives (e.g., acceptance or rejection) and result in corresponding action and subsequent

observable behavioral patterns of implicit motives (see Fyans, Salili, Maehr, & Desai, 1983; Maehr & Nicholls, 1980). In this context, the conduction of experimental motive arousal studies that played an important role in determining the original motive coding systems (e.g., McClelland, Atkinson, Clark, & Lowell, 1953; Winter, 1973) may also help to determine under which conditions motives tend to be elicited in various cultures. Such an approach could also contribute to an examination of category boundaries of the motive scoring systems across cultures.

To conclude, many cross-cultural studies on implicit motives with non-Western cultures only conceptualized culture-inherent stimuli. Therefore, results cannot directly be compared across cultures, and cross-cultural commonalities can hardly be detected. This study demonstrates that the extended methodological cross-cultural analysis led to unbiased, culture-independent sets of picture stimuli that can be used for cross-cultural comparisons of assessed data on two implicit motives, nAffiliation and nPower, respectively. Given the unique contribution that implicit variables may add to the cross-cultural study of personality, this seems to be of great value for motivational research. In a next step, findings on motive strength should be examined carefully. In this context, one should aim to identify common criteria of external validation of implicit motives that would represent further evidence for the meaningfulness of our findings. Even though the presented procedure may be complex and time-consuming, derived findings seem to be promising for future studies that may focus on the need for achievement that was not addressed in the present article in a similar way. Therefore, a good knowledge of participants' cultural background and a close collaboration with local experts is necessary to apply appropriate methods for data assessment and standards of data interpretation.

NOTES

1. Winter's (1991) scoring system combines nIntimacy and nAffiliation into a single image category because of their theoretical and empirical overlap. Therefore, the term *nAffiliation* will be used to refer to the combined intimacy-affiliation motive in the following.

2. We did not focus on need for achievement because this methodological study was planned as a first step in cross-cultural research on interrelations of affiliation motivation, power motivation, and reproductive indicators derived from an individual's life history.

3. Analyses based on only the first striving given by the participants for both descriptions showed similar results—that is, percentages of the responses assigned to scoring categories do not sufficiently differ.

4. The remaining answers of the students (advice: $n = 124$; 25.2%; benefit: $n = 170$; 33.5%) could not be assigned to one of the three striving categories. Among those not-categorized responses, mostly religious reasons for an individual's behavior were reported by the participants.

5. In Cameroon, we did not train study assistants to code picture stories because no experienced students of psychology were at hand. Also, no language barrier emerged (English stories). However, we discussed coding of problematic stories with our colleague Relindis Yovsi at the University of Osnabrück, who is ethnic Nso.

6. Percentage agreement between scorers across all stories was conservatively estimated by the index of concordance: $2 \times$ number of agreements between scorers / (scorer A scores + scorer B scores) (see Martin & Bateson, 1993; Winter, 1991).

7. Bias analyses using loglinear models confirmed findings derived by the Mantel-Haenszel procedure by pointing to the same four unbiased picture cards.

8. As the value type hedonism overlaps with both achievement and stimulation, it was not considered in analyses on cross-cultural differences of higher-order dimensions.

9. Dichotomization of motive scores did not negatively affect the meaningfulness of our data. Generally, we found high correlations between the two types of scores ($r = .73$ to .95; p's $< .01$) within our study samples (pretests and main study). Also, no significant differences were found when comparing correlations between total motive strength and both, nondichotomized and dichotomized scores by employing Fisher's r-to-z transformation.

10. By using likelihood ratio comparisons (e.g., Tabachnik & Fidell, 1996) between the nested models examining nonuniform and uniform bias, respectively, it was shown for each of the final picture cards that the most parsimonious model, which only included score-level group as a factor, adequately fits the data for the two given motives. For the total sample, the averaged interitem correlations (Cronbach's alpha) are .37 for nAffiliation and .28 for nPower, respectively.

11. The ability index was calculated on the basis of data assessed with the culturally bias-free, four-picture sets.

12. In our opinion, such analyses based on context variables range in the intersection of detection of method and item bias.

REFERENCES

Allen, J., & Walsh, J. A. (2000). A construct-based approach to equivalence: Methodologies for cross-cultural/multicultural personality assessment research. In R. H. Dana (Ed.), *Handbook of cross-cultural and multicultural personality assessment* (pp. 63-85). Mahwah, NJ: Lawrence Erlbaum.

Biernat, M. (1989). Motives and values to achieve: Different constructs with different effects. *Journal of Personality*, *57*, 69-95.

Bilsky, W., & Peters, M. (1999). Estructura de los valores y la religiosidad. Una investigación comparada realizada en México [The structure of values and religiousness. A comparative study in Mexico]. *Revista Mexicana de Psicologia*, *16*, 77-88.

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2002). Different kinds of DIF: A distinction between absolute and relative forms of measurement invariance and bias. *Applied Psychological Measurement*, *26*, 433-450.

Brislin, R. W. (1993). *Understanding culture's influence on behavior.* New York: Harcourt Brace.

Camilli, G., & Shepard, L. A. (1994). *Methods for identifying biased test items.* Thousands Oaks, CA: Sage.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81-105.

Campos, D. (1991). *La personalidad autoritaria y participación política* [The authoritarian personality and political participation]. San José, Costa Rica: Cuadernos del CSUCA.

Church, A. T., & Katigbak, M. S. (1992). The cultural context of academic motives: A comparison of Filipino and American college students. *Journal of Cross-Cultural Psychology*, *23*, 40-58.

Costantino, G., Malgady, R. G., & Rogler, L. H. (1988). *TEMAS (Tell-me-a-story) manual*. Los Angeles: Western Psychological Services.

Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), *Educational measurement* (pp. 433-507). Washington, DC: American Council on Education.

Dana, R. H. (1962). The validation of projective techniques. *Journal of Projective Techniques*, *26*, 182-186.

DeCharms, R. (1992). Personal causation and the origin concept. In C. P. Smith (Ed.), *Motivation and personality: Handbook of thematic content analysis* (pp. 325-333). New York: Cambridge University Press.

De Vos, G. A. (1968). Achievement and innovation in culture and personality. In E. Norbeck, D. Price-Williams, & W. M. McCord (Eds.), *The study of personality: An interdisciplinary approach* (pp. 348-370). New York: Holt, Rinehart & Winston.

Dodeen, H. (2004). Stability of differential item functioning over a single population in survey data. *Journal of Experimental Education*, *72*, 181-193.

Doi, K. (1982). A two dimensional theory of achievement motivation: Affiliative and non-affiliative. *Japanese Journal of Psychology*, *52*, 344-350.

Dorans, N. J., & Kulick, E. (1986). Demonstrating the utility of the standardization approach to assessing unexpected differential item performance on the Scholastic Aptitude Test. *Journal of Educational Measurement*, *23*, 355-368.

Eckensberger, L. H. (2002). Paradigms revisited: From incommensurability to respected complementarity. In H. Keller, Y. Poortinga, & A. Schölmerich (Eds.), *Between culture and biology: Perspectives on ontogenetic development integrating diverse perspectives* (pp. 341-383) Cambridge, UK: Cambridge University Press.

Emmons, R. A. (1999). *The psychology of ultimate concerns: Motivation and spirituality in personality.* New York: Guilford.

Emmons, R. A., & McAdams, D. P. (1991). Personal strivings and motive dispositions: Exploring the links. *Personality and Social Psychology Bulletin*, *17*, 648-654.

Engelhard, G., Hansche, L., & Rutledge, K. E. (1990). Accuracy of bias review judges in identifying differential item functioning on teacher certification tests. *Applied Measurement in Education*, *3*, 347-360.

Entwisle, D. R. (1972). To dispel fantasies about fantasy-based measures of achievement motivation. *Psychological Bulletin*, *77*, 377-391.

Ervin, S. M. (1964). Language and TAT content in French-English bilinguals. *Journal of Abnormal and Social Psychology*, *68*, 500-507.

Espe, H. (1985). A cross-cultural investigation of the graphic differential. *Journal of Psycholinguistic Research*, *14*, 97-111.

French, E. G., & Lesser, G. S. (1964). Some characteristics of the achievement motive in women. *Journal of Abnormal and Social Psychology*, *68*, 119-128.

Fyans, L. J., Jr., Salili, F., Maehr, M. L., & Desai, K. A. (1983). A cross-cultural exploration into meaning of achievement. *Journal of Personality and Social Psychology*, *44*, 1000-1013.

Hagger, M. S., Biddle, S. J. H., Chow, E. W., Stambulova, N., & Kavussanu, M. (2003). Physical self-perceptions in adolescence: Generalizability of a hierarchical multidimensional model across three cultures. *Journal of Cross-Cultural Psychology*, *34*, 611-628.

Heyns, R. W., Veroff, J., & Atkinson, J. W. (1958). A scoring manual for the affiliation motive. In J. W. Atkinson (Ed.), *Motives in fantasy, action and society* (pp. 205-218). Princeton, NJ: Van Nostrand.

Hofer, J., & Chasiotis, A. (2003). Congruence of life goals and implicit motives as predictors of life satisfaction: Cross-cultural implications of a study of Zambian male adolescents. *Motivation and Emotion*, *27*, 251-272.

Hofer, J., & Chasiotis, A. (2004). Methodological considerations of applying a TAT-type picture-story-test in cross-cultural research: A comparison of German and Zambian adolescents. *Journal of Cross-Cultural Psychology*, *35*, 224-241.

Hofstede, G. (2000). *Culture's consequences. International differences in work-related values*. Thousand Oaks, CA: Sage.

Holland, P. W., & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer & H. I. Braun (Eds.), *Test validity* (pp. 129-145). Hillsdale, NJ: Lawrence Erlbaum.

Holland, P. W., & Wainer, H. (1993). *Differential item functioning*. Hillsdale, NJ: Lawrence Erlbaum.

Holtzman, W. H. (1980). Projective techniques. In H. C. Triandis & R. W. Brislin (Eds.), *Handbook of cross-cultural psychology* (Vol. 2, pp. 245-278). Boston: Allyn & Bacon.

Huber, H. (1989). Sozial- und kulturanthropologische Theorien zur Sozialisationsforschung [Social- and cultural-anthropological theories of research on socialization]. In G. Trommsdorff (Ed.), *Sozialisation im Kulturvergleich* [Socialization in cross-cultural perspective] (pp. 25-41). Stuttgart, Germany: Enke.

Kagan, S., & Knight, G. P. (1981). Social motives among Anglo American and Mexican American children: Experimental and projective measures. *Journal of Research in Personality*, *15*, 93-106.

Keller, H., & Greenfield, P. M. (2000). History and future development in cross-cultural psychology. *Journal of Cross-Cultural Psychology*, *31*, 52-62.

Klinger, E. (1966). Fantasy need achievement as a motivational construct. *Psychological Bulletin*, *66*, 291-308.

Klinger, E., & McNelly, F. W., Jr. (1969). Fantasy need achievement and performance: A role analysis. *Psychological Review*, *76*, 574-591.

Kok, F. K., Mellenbergh, G. J., & Van Der Flier, H. (1985). Detecting experimentally induced item bias using the iterative logit method. *Journal of Educational Measurement*, *22*, 295-303.

Kornadt, H.-J., Eckensberger, L. H., & Emminghaus, W. B. (1980). Cross-cultural research on motivation and its contribution to a general theory of motivation. In H. C. Triandis & R. W. Brislin (Eds.), *Handbook of cross-cultural psychology* (Vol. 3, pp. 223-321). Boston: Allyn & Bacon.

Kornadt, H.-J., & Voight, E. (1970). *Situation und Entwicklungsprobleme des Schulsystems in Kenia* [Situation and developmental problems of the school system in Kenya]. Stuttgart, Germany: Ernst Klett Verlag.

Lee, S. G. (1953). *Manual of a Thematic Apperception Test for African subjects*. Pietermaritzburg, South Africa: University of Natal Press.

Lindzey, G. (1961). *Projective techniques and cross-cultural research*. New York: Appleton-Century-Crofts.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum.

Lundy, A. (1985). The reliability of the Thematic Apperception Test. *Journal of Research in Personality*, *49*, 141-145.

Maehr, M. L. (1974). Culture and achievement motivation. *American Psychologist*, *29*, 887-896.

Maehr, M. L., & Nicholls, J. G. (1980). Culture and achievement motivation: A second look. In N. Warren (Ed.), *Studies in cross-cultural psychology* (Vol. 2, pp. 221-267). New York: Academic Press.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*, 224-253.

Martin, P., & Bateson, P. (1993). *Measuring behavior: An introductory guide*. New York: Cambridge University Press.

McAdams, D. P. (1980). A thematic coding system for the intimacy motive. *Journal of Research in Personality*, *14*, 413-432.

McAdams, D. P. (1992). The intimacy motive. In C. P. Smith (Ed.), *Motivation and personality: Handbook of thematic content analysis* (pp. 224-228). New York: Cambridge University Press.

McAdams, D. P., & Vaillant, G. E. (1982). Intimacy, motivation, and psycho-social adjustment: A longitudinal study. *Journal of Personality Assessment*, *46*, 586-593.

McClelland, D. C. (1961). *The achieving society*. Princeton, NJ: Van Nostrand.

McClelland, D. C. (1975). *Power: The inner experience*. New York: Irvington.

McClelland, D. C. (1985). *Human motivation*. New York: Cambridge University Press.

McClelland, D. C. (1987). Characteristics of successful entrepreneurs. *Journal of Creative Behavior*, *3*, 219-233.

McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. (1953). *The achievement motive*. New York: Appleton-Century-Crofts.

McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review*, *96*, 690-702.

McClelland, D. C., & Pilon, D. A. (1983). Sources of adult motives in patterns of parent behavior in early childhood. *Journal of Personality and Social Psychology*, *44*, 564-574.

McClelland, D. C., & Winter, D. G. (1969). *Motivating economic achievement*. New York: Free Press.

Mellenbergh, G. J. (1982). Contingency table models for assessing item bias. *Journal of Educational Statistics*, *7*, 105-118.

Mellenbergh, G. J. (1989). Item bias and item response theory. *International Journal of Educational Research*, *13*, 127-143.

Meyer, J. M., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., et al. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist*, *56*, 128-165.

Millsap, R. E., & Everson, H. T. (1993). Methodology review: Statistical approaches for assessing bias. *Applied Psychological Measurement*, *17*, 297-334.

Murray, H. A. (1943). *Thematic Apperception Test manual*. Cambridge, MA: Harvard University Press.

Murstein, B. I. (1965a). Reliability. In B. I. Murstein (Ed.), *Handbook of projective techniques* (pp. 189-218). New York: Basic Books.

Murstein, B. I. (1965b). The stimulus. In B. I. Murstein (Ed.), *Handbook of projective techniques* (pp. 509-546). New York: Basic Books.

Nohlen, D., & Nuscheler, F. (1993). *Handbuch der Dritten Welt, Band 1: Grundprobleme Theorien Strategien* [Third world handbook, Vol. 1: Basic problems, theories, and strategies]. Bonn, Germany: J. H. W. Dietz Nachf. GmbH.

Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill.

Oishi, S., Schimmack U., Diener, E., & Suh, E. (1998). The measurement of values and individualism-collectivism. *Personality and Social Psychology Bulletin*, *24*, 1177-1189.

Okazaki, S., & Sue, S. (1995). Methodological issues in assessment research with ethnic minorities. *Psychological Assessment*, *7*, 367-375.

Peterson, B. E., & Stewart, A. J. (1993). Generativity and social motives in young adults. *Journal of Personality and Social Psychology*, *65*, 186-198.

Plake, B. S. (1980). A comparison of a statistical and subjective procedure to ascertain item validity: One step in the test validation process. *Educational and Psychological Measurement*, *40*, 397-404.

Poortinga, Y. H. (1989). Equivalence of cross-cultural data: An overview of basic issues. *International Journal of Psychology*, *24*, 737-756.

Potenza, M. T., & Dornas, N. J. (1995). DIF assessment for polytomously scored items: A framework for classification and evaluation. *Applied Psychological Measurement*, *19*, 23-37.

Realo, A., Koido, K., Ceulemans, E., & Allik, J. (2002). Three components of individualism. *European Journal of Personality*, *16*, 163-184.

Rudner, L. M., Getson, P. R., & Knight, D. L. (1980). A Monte Carlo comparison of seven biased item detection techniques. *Journal of Educational Measurement*, *17*, 1-10.

Schultheiss, O. C. (2001). An information processing account of implicit motive arousal. In M. L. Maehr & P. R. Pintrich (Eds.), *Advances in motivation and achievement: Vol. 12. New directions in measures and methods* (pp. 1-41). Greenwich, CT: JAI Press.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 1-65). Orlando, FL: Academic Press.

Schwartz, S. H. (1994a). Are there universal aspects in the structure and contents of human values? *Journal of Social Issues*, *50*, 19-45.

Schwartz, S. H. (1994b). Beyond individualism/collectivism: New cultural dimensions of values. In U. Kim, H. C. Triandis, C. Kagitcibasi, S.-C. Choi, & G. Yoon (Eds.), *Individualism and collectivism: Theory, method, and applications* (pp. 85-119). Thousands Oaks, CA: Sage.

Schwartz, S. H., & Bilsky, W. (1990). Toward a theory of the universal content and structure of values: Extensions and cross-cultural replications. *Journal of Personality and Social Psychology*, *58*, 878-891.

Schwartz, S. H., Melech, G., Lehmann, A., Burgess, S., Harris, M., & Owen, V. (2001). Extending the cross-cultural validity of the theory of basic human values with a different method of measurement. *Journal of Cross-Cultural Psychology*, *32*, 519-542.

Schwartz, S. H., & Sagiv, L. (1995). Identifying culture-specifics in the content and structure of values. *Journal of Cross-Cultural Psychology*, *26*, 92-116.

Serpell, R. (1993). *The significance of schooling: Life-journeys in an African society*. Cambridge, UK: Cambridge University Press.

Shimahara, N. (1970). Enculturation: A reconsideration. *Current Anthropology*, *11*, 143-154.

Singelis, T. M., Triandis, H. C., Bhawuk, D. P. S., & Gelfand, M. J. (1995). Horizontal and vertical dimensions of individualism and collectivism: A theoretical and measurement refinement. *Cross-Cultural Research*, *29*, 240-275.

Smith, C. P. (Ed.) (1992). *Motivation and personality: Handbook of thematic content analysis*. New York: Cambridge University Press.

Smith, C. P., Feld, S. C., & Franz, C. E. (1992). Methodological considerations: Steps in research employing content analysis systems. In C. P. Smith (Ed.), *Motivation and personality: Handbook of thematic content analysis* (pp. 515-536). New York: Cambridge University Press.

Spangler, W. D. (1992). Validity of questionnaire and TAT measures of need for achievement: Two meta-analyses. *Psychological Bulletin*, *112*, 140-154.

Steenkamp, J.-B. E. M., & Baumgartner, H. (1981). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, *25*, 78-90.

Super, C. M. (1983). Cultural variation in the meaning and uses of children's "intelligence." In J. B. Deregowski, S. Dziurawiec, & R. C. Annis (Eds.), *Expiscations in cross-cultural psychology* (pp. 199-212). Lisse, the Netherlands: Swets & Zeitlinger.

Swamanithan, H., & Rogers, H. J. (1990). Detecting differential item functioning using the logistic regression procedure. *Journal of Educational Measurement*, *27*, 361-370.

Tabachnik, B. G., & Fidell, L. S. (1996). *Using multivariate statistics*. New York: HarperCollins.

Triandis, H. C. (1995). *Individualism and collectivism*. Boulder, CO: Westview.

Triandis, H. C. (1996). The psychological measurement of cultural syndromes. *American Psychologist*, *51*, 407-415.

Tuerlinckx, F., De Boeck, P., & Lens, W. (2002). Measuring needs with the Thematic Apperception Test: A psychometric study. *Journal of Personality and Social Psychology*, *82*, 448-461.

United Nations Development Programme. (2003). *Human development report 2003*. New York: Oxford University Press.

Van Der Flier, H., Mellenbergh, G. J., Adèr, H. J., & Wijn, M. (1984). An iterative item bias detection method. *Journal of Educational Measurement*, *21*, 131-145.

Van de Vijver, F. J. R. (1994). Bias: Where psychology and methodology meet. In A.-M. Bouvy, F. J. R. van de Vijver, P. Boski, & P. Schmitz (Eds.), *Journeys into cross-cultural psychology* (pp. 111-126). Lisse, the Netherlands: Swets & Zeitlinger.

Van de Vijver, F. J. R. (2000). The nature of bias. In R. H. Dana (Ed.), *Handbook of cross-cultural and multicultural personality assessment* (pp. 87-106). Mahwah, NJ: Lawrence Erlbaum.

Van de Vijver, F. J. R., & Leung, K. (1997). *Methods and data analysis for cross-cultural research*. Thousands Oaks, CA: Sage.

Van de Vijver, F. J. R., & Poortinga, Y. H. (1997). Towards an integrated analysis of bias in cross-cultural assessment. *European Journal of Psychological Assessment*, *13*, 29-37.

Van Leest, P. F. (1997). Bias research in the Netherlands. *European Review of Applied Psychology*, *47*, 319-327.

Veroff, J. (1992). Thematic apperceptive methods in survey research. In C. P. Smith (Ed.), *Motivation and personality: Handbook of thematic content analysis* (pp. 100-109). New York: Cambridge University Press.

Veroff, J., Feld, S., & Crockett, H. (1966). Explorations into the effects of picture cues on thematic apperceptive expression of achievement motivation. *Journal of Personality and Social Psychology*, *3*, 171-181.

Wainer, H., & Braun, H. I. (1988). *Test validity*. Hillsdale, NJ: Lawrence Erlbaum.

Weinberger, J., & McClelland, D. C. (1990). Cognitive versus traditional motivational models. Irreconcilable or complementary? In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition. Foundations of social behavior* (Vol. 2, pp. 562-597). New York: Guilford.

Weinstein, M. S. (1969). Achievement motivation and risk preference. *Journal of Personality and Social Psychology*, *13*, 153-172.

Welkenhuysen-Gybels, J. (2003). *The detection of differential item functioning in Likert score items*. Unpublished doctoral dissertation, Catholic University of Leuven, Belgium.

Welkenhuysen-Gybels, J., & Billiet, J. (2002). A comparison of techniques for detecting cross-cultural inequivalence at the item level. *Quality and Quantity*, *35*, 197-218.

Winter, D. G. (1973). *The power motive*. New York: Free Press.

Winter, D. G. (1991). *Manual for scoring motive imagery in running text*. Ann Arbor: University of Michigan.

Winter, D. G. (1992). Power motivation revisited. In C. P. Smith (Ed.), *Motivation and personality: Handbook of thematic content analysis* (pp. 301-310). New York: Cambridge University Press.

Winter, D. G., & Stewart, A. J. (1977). Power motive reliability as a function of retest instructions. *Journal of Consulting and Clinical Psychology*, *45*, 436-440.

Yovsi, R. D. (2003). *Ethnotheories about breastfeeding and mother-infant interaction*. Münster, Germany: LIT Verlag.

Yu, A. (1996). Ultimate life concerns, self, and Chinese achievement motivation. In M. H. Bond (Ed.), *The handbook of Chinese psychology* (pp. 227-246). Hong Kong: Oxford University Press.

Jan Hofer received his Ph.D. from the Friedrich-Alexander University, Erlangen-Nuremberg. He is currently a member of the research group Cross-Cultural Life-Span Psychology at the University of Osnabrück. His research interests include methodological problems in cross-cultural research, the development and relationship of implicit and explicit motives, self-development, and correlates of subjective well-being.

Athanasios Chasiotis received his Ph.D. from the University of Osnabrück and is currently head of the research group Cross-Cultural Life-Span Psychology at this university. His research interests concern the intersection of evolutionary and cross-cultural developmental psychology. He is co-editor (with E. Voland and W. Schiefenhövel) of the forthcoming book Grandmotherhood—The Evolutionary Significance of the Second Half of Life *(Rutgers University Press).*

Wolfgang Friedlmeier received his Ph.D. in psychology from the University of Konstanz, Germany (1993). He is currently assistant professor for cross-cultural psychology at Grand Valley State University in Allendale, MI. His main research areas are cross-cultural psychology and its methodology, socioemotional development, and socialization. He is co-editor (with P. Chakkarath and B. Schwarz) of the book Culture and Human Development *(2005, Psychology Press) and co-author (with M. Holodynski) of the book* Development of Emotions and Emotion Regulation *(2005, Springer).*

Holger Busch is currently working on his diploma thesis concerning generativity in various cultures at the University of Osnabrück. Working as student assistant for the research group Cross-Cultural Life-Span Psychology, he has repeatedly been on field stays in Cameroon. His research interests are human motivation and generativity in diverse cultural contexts.

Domingo Campos is full professor of psychology and director of the Psychological Research Institute of the University of Costa Rica. His main research interests are in the areas of children's cognitive and social development and behavioral genetics.