An Examination of Acquiescent Response Styles in Cross-Cultural Research

Ronald Fischer, *Centre for Applied Cross-Cultural Research, Wellington, New Zealand*, Ronald.Fischer@vuw.ac.nz

Johnny R. J. Fontaine, *Gent University, Belgium*

Fons J. R. van de Vijver, *Tilburg University, The Netherlands, and North-West University, South Africa*

Dianne A. van Hemert, *University of Amsterdam, The Netherlands*

Response styles constitute a formidable challenge for cross-cultural research. In this article, three different response styles are discussed (acquiescence, extremity scoring, and social desirability). Acquiescence responding (*ARS*) is then integrated into a larger classical test theoretical framework, which allows for an examination of the various roles that *ARS* may play in cross-cultural research. A new meta-analytical method is proposed to examine the prevalence and nature of *ARS*. Preliminary evidence suggests that *ARS* has only a small, but systematic effect on survey responses. The meaning of *ARS* is explored through correlations with nation-level indicators. Implications for future research are discussed.

What is Style and What is Bias in Cross-Cultural Comparisons?

Response styles have long been identified as a major threat to survey research (e.g., Guilford, 1954). This problem may be aggravated in cross-cultural research since individuals of different cultural backgrounds may use answer-response scales in different ways (Smith, 2004). Therefore, the aim of the current manuscript is to investigate to what extent response styles are operating in a cross-cultural context and how they can be isolated. Three types of response styles are discussed in the next section (acquiescence, extremity scoring, and social desirability). Then, the literature on a specific and commonly discussed type of response style (acquiescence) is being reviewed. Finally, one possible approach to investigate direct effects of acquiescence across cultures is proposed and preliminary data are presented that investigate the prevalence and meaning of *ARS*.

Three Common Types of Response Styles

Three major types of response styles have been discussed in the literature. The first is acquiescent response style (*ARS*), which is the tendency to either agree or disagree independent of item content (Cheung & Rensvold, 2000). *ARS* usually leads to a shift of the mean of the item in an upward or downward direction. The second type is extreme response style (ERS), which is the tendency to use either (only) moderate or extreme categories of rating scales. Irrespective of item content, individuals either agree or disagree with an item content strongly or they tend to use only the middle categories (modesty style). The third response style often mentioned in the literature is social desirability responding (SDR; Crowne & Marlowe, 1960). This is a tendency to respond in a socially desirable way, which amounts to responding in a way that is expected to get approval by significant others of the respondent. Paulhus (1991) distinguished between impression management and self-deception. Impression management is a conscious strategy to appear or present oneself in a positive light, whereas self-deception is an unconscious tendency to see and portray oneself in a socially acceptable way. The notion of self-deception introduces an important new element in the conceptualization of response styles. These styles may reflect important personality features that cannot be simply dismissed as

measurement disturbances that should be eliminated (as seems to be the conventional wisdom, e.g. Guilford, 1954). It has been argued that social desirability is an aspect of Agreeableness (McCrae & Costa, 1983). Conceptually, the three styles are distinct, even though in particular *ARS* and SDR may seem to refer to related, if not identical constructs. However, SDR is by definition directly linked to the item content (Paulhus, 1984), whereas *ARS* is the tendency to use certain response categories independent of item content (Cheung & Rensvold, 2000). The three styles should also be assessed and analyzed independently. However, if there are only positively or negatively phrased items (all items are scaled in the same direction), ERS and *ARS* are confounded and cannot be separated (Baumgartner & Steenkamp, 2001; Chun, Campbell & Yoo, 1974).

Acquiescent Responding

Acquiescent Responding (*ARS*) has to be viewed as method bias that has a systematic impact on scores and can affect the nature and size of cross-cultural differences. Previous research has found that individuals' *ARS* was negatively associated with socioeconomic status, level of education, and acculturation status, and positively with age; furthermore, some studies found ethnic differences (Bachman & O'Malley, 1984; Carr & Krause, 1978; Greenleaf, 1992; Marin, Gamba, & Marin, 1992; Phillips & Clancy, 1970; Ross & Mirowsky, 1984; Winkler, Kanouse, & Ware, 1982; Yu & Murphy, 1993). These differences have been explained in different ways, countries scoring higher on *ARS* were looser (in terms of norms and restrictions), more collectivistic, more feminine, lower in power distance and lower in Gross Domestic Product (GDP), while for uncertainty avoidance both a positive and a negative association has been found (Boldt, 1976; Johnson, Kulesa, Cho, & Shavitt, 2005; Smith, 2004; Smith & Fischer 2007; Van Herk, Poortinga, & Verhallen, 2004). These findings suggest that there is systematic variability in *ARS* at both individual and culture level. Observed scores are likely to be systematically influenced by response tendencies reflecting both individual difference variables (*IARS*, individual response styles) as well as cultural processes (*CARS*, cultural response styles).

The aforementioned discussion also suggests that *ARS* (overall *ARS*, independent of the source of variation) is not uniform but varies with method factors. Studies by Greenleaf (1992) and Baumgartner and Steenkamp (2001) suggested that *ARS* might be method dependent. Ray (1983) argued that *ARS* increases with ambiguity of the scale. These findings imply that there may be interactions between response tendencies and particular method variables. Acquiescence may increase if the number of response options does not allow individuals to express themselves adequately (Hui & Triandis, 1985, 1989). Particular scales or items might be more ambiguous than others. Therefore, *ARS* is likely to interact with specific method variables. Also, *ARS* may vary with content-related variables and may be stronger when assessing more personal domains (Van Dijk, Datema, Piggen, Welten, & Van de Vijver, this volume).

Investigating *ARS*

The classical test approach implies that trait scores (or true underlying scores) should capture most variance in observed scores, but there are likely to be consistent influences of method variables, which include response tendencies. Since *ARS* as part of a method component is supposed to be systematic rather than random, these effects could be investigated. Such an endeavor should focus on a variety of different and preferably uncorrelated theoretical variables. An important issue is the separation of *ARS* and substantive meaning. Examining theoretically related variables (in the absence of valid external criteria as is often the case in survey research), it is often difficult to evaluate the relative contribution of substantive factors and response styles to observed responses (Baumgartner & Steenkamp, 2001). Thus, examining

theoretically unrelated variables would serve the purposes of the present study since they would avoid confounding of substantive information and response style effects.

Collecting new data to isolate these effects is time-intensive and requires substantial resources. An alternative and powerful tool could be to meta-analytically review and summarize existing research. The main effects of *ARS* could be identified through an investigation of the mean across a large number of content-independent items or scales that have positively and negatively phrased items (Baumgartner & Steenkamp, 2001). Using instrument-based meta-analyses (Van Hemert, Van de Vijver, Poortinga, & Georgas, 2002; Van Hemert, Van de Vijver, & Poortinga, 2002), indicators for a range of unrelated constructs across a number of cultural groups could be derived. The sample means are first converted to a common metric (i.e., means are standardized to have a range of 0-1, Fischer & Chalmers, 2008). The standardized means are then averaged per culture (using weighting formulas provided by Lipsey & Wilson, 2001). This will provide an overall indicator of *ARS* per country (called here meta-analytic *ARS* index or *MARS*). The aggregated means across areas and instruments at a cultural group level could then be analyzed for patterns indicating uniform and consistent response styles (using correlational techniques, cluster analysis, factor analysis, multi-dimensional scaling analysis, variance-decompositioning approaches; only the correlational approach results are reported in this chapter). Assuming *MARS* is based on a range of theoretically uncorrelated constructs measured in independent samples it can be seen as a true estimator of *ARS* and will allow us to examine whether there are uniform and consistent effects that threaten validity of survey research. Such a meta-analytical analysis by necessity would be conducted at the nation level and give only an estimation of *CARS* (*cultural response styles*). The impact of *CARS* on observed scores is especially important for cross-cultural ecological nation-level research (Hofstede, 1980) and the analysis would indicate to what extent such analyses are influenced by *CARS*.

It would be possible to derive similar indicators at the individual level. Re-analyses of existing large datasets could be used to simultaneously find indicators for *IARS* (*individual response styles*). Therefore, two complementary analyses at the cultural level using meta-analysis and at the individual level using re-analyses of existing datasets can be used to estimate the extent of *ARS* at the two levels. This analysis would show to what extent *ARS* is systematic and stable at each level.

Second, meta-analysis could be used to investigate the influence on *CARS* of the type of scale being used, number of response options or ambiguity of items. Similar analyses of *IARS* could be conducted through re-analyses of existing data sets. Large cross-cultural data sets also enable the study of *IARS* and *CARS* in the same data set. Thus, Smith and Fischer (2008) investigated interactions between *IARS* and *CARS* in a large data set using organizational data from various cultures.

Table 1. Variables Influencing Response Styles

	Individual level	Nation level
Psychological variables	Personality	Norms and values
Sociological/economical context variables	Socio-economic status, education	Macro-economic (affluence) and political indicators

Finally, as stated before, we investigate to what extent *ARS* is stable at both levels. The current assumption in the literature is that *ARS* is a specific type of a systematic method component. As a consequence, it would be possible to examine the sources of variation underlying *ARS* at the individual and cultural level. Our literature review above indicates that both psychological and contextual variables might be associated with response tendencies. Education and income are contextual, non-psychological variables, whereas personality and

values are psychological variables (Table 1). Using the *IARS* and *CARS (MARS)* indicators described above, it would be possible to test relationships between response tendencies and contextual variables at the two levels empirically. The overall response style indicators can be correlated with cultural as well as economic, religious, education, population, mass communication and ecology indices at the nation-level. However, it is important to use contextual indicators for these analyses that are not based on rating scales since such an approach would confound the method (ratings of contextual variables) with the construct of interest (acquiescent responding when answering rating scales). In the next part we will demonstrate our meta-analytic approach by presenting some preliminary data on the main effect of *CARS* and explore some potential meanings of *CARS*.

Method

Variables recorded in meta-analysis

    Separate meta-analyses for the following scales and instruments were conducted: Stait-Trait Anxiety Inventory (Spielberger, Gorsuch & Lushene, 1970), General Health Questionnaire (Goldberg, 1972), Achievement Goal inventories (Elliot & Church, 1997; Middleton & Midgley, 1997), Self-Esteem (Rosenberg, 1965, 1979), Organizational Commitment (various scales, Fischer & Mansell, in press), Organizational Justice (various scales, Fischer, 2008) and Transformational Leadership (Bass, 1985). These scales are used in a variety of disciplines (clinical, educational, health, organizational, social and personality psychology) and capture self-ratings (anxiety, general health, achievement goals, and self-esteem), self-ratings in relation to some referent (e.g., commitment to one's organization) and ratings of others and external entities (organizational justice, transformational leadership). The latter type of ratings have been shown to be less affected by common method variance (Crampton & Wagner, 1994), increasing the validity of our approach. Literature search to retrieve relevant studies were conducted in PsycINFO, Social Sciences Citation Index; other relevant electronic sources and reviews were also consulted. The data presented here are preliminary since coding is continuing (for non-English articles on GHQ, self-esteem, STAI, transformational leadership; and for articles published prior to 1990 on self-esteem and GHQ). The number of participants and countries for which data are available for each instrument and its subscales (including brief definitions of the scales) are reported in Table 2.

    We recorded the mean or sum reported in each individual study. Since these instruments used different response scales (typically Likert format ranging from agree - disagree format to 100 point scales) and different numbers of items, we standardized each mean or sum, by dividing calculated means by the number of response options. Therefore, all indicators have a possible range from 0 to 1 (Fischer & Chalmers, 2008).

Other variables used for analyses

    Indicators of acquiescence responding. A number of large-scale studies have reported indicators of response styles. First, Smith et al. (2002) reported the mean reliance of managers on a number of different sources of guidance across situations. Higher scores are thought to capture greater acquiescence (Smith, 2004). Mean value ratings across all 57 values in the Schwartz Value Survey (SVS; Schwartz, 1992) are available from Schwartz (2004) and Spini (2001). Finally, McCrae, Terracciano and 78 members of the Personality Profiles of Culture Project (2005) reported acquiescence scores based on observer ratings of personality dimensions across a large number of countries and targets.

    Control variable: We used the quality index provided by McCrae et al. (2005), which is the mean of a number of quality-related indicators of a large personality study (number of missing information on response forms, number of individuals not discriminating between items, translation adequacy, etc.).

Ecological and economic indicators were derived from Georgas and Berry (1987) and Van Hemert et al. (2002b). Indicators included average annual growth rate of GDP between 1980 and 1987; GDP in 1987; population density in 1987; income inequality in 1999 (as measured by the Gini index); Purchasing Power Parity in 1997 and Human Development Index in 1987. These variables are indicators of socioeconomic and contextual variables at the nation-level (Table 1) and are not derived from rating scales.

Table 2. Descriptive Information on Available Measures

Instrument & Subscales	Definition	General Sample Characteristics	Participants & countries (available data)	
STAI-Trait Anxiety	Habitual or constant anxiety levels	General population and clinical control groups	26,402	28
STAI-State Anxiety	Transient or temporal anxiety levels	General population and clinical control groups	19,844	30
GHQ	General mental health	General population and clinical control groups	76,852	21
Achievement motivation goals Performance approach goals	Orientation to display or prove ability and outdoing others	Adolescents (primarily high school)	32,264	13
Achievement motivation goals Mastery goals	Orientation towards learning new things, improving ability or understanding, and mastering the material or task	Adolescents (primarily high school)	29,372	12
Achievement motivation goals Performance avoidance goals	Goals aimed at avoiding failure, protecting oneself from being embarrassed, and being judged by others as lacking ability and competence	Adolescents (primarily high school)	22,352	13
Self-esteem	General self-esteem	General population and student samples	40,086	24
Organizational commitment Affective commitment	Identification with and internalization of the goals and values of the organization	Working adults	105,335	49
Organizational commitment Continuance commitment	Calculation of costs and benefits associated with staying in the organization	Working adults	15,734	30
Organizational commitment Normative commitment	Feeling obliged to stay because of social pressures	Working adults	12,204	30
Organizational Justice Distributive justice	Evaluation of the rewards received	Working adults	30,528	29
Organizational Justice Procedural justice	Evaluation of the formal procedures used during organizational decision-making	Working adults	76,367	29
Organizational Justice Interactional Justice	Evaluation of the enactment of procedures by one's supervisor	Working adults	67,060	23
Transformational Leadership (MLQ)	Assessment of leadership behavior	Working adults	20,236	14
Schwartz Value Survey (Schwartz, 2005)	Mean rating across 55 human values	Students, Teachers	41,968	69
Schwartz Value Survey (Spini, 2001)	Mean rating across 55 human values	Students	3,787	21
NEO-Acquiescence Other Rating (McCrae et al., 2005)	Acquiescence rating of personality observer ratings	Students	12,156	51
Sources of Guidance (Smith et al., 2002)	Acquiescence rating across all sources of guidance for business leaders	Business leaders	7,091	47

We also used a range of <u>political rights and democracy-related indicators</u> (taken from Van Hemert et al., 2002b). These indicators were derived from work by Humana (1986, Human Rights index for rights and freedoms), Freedom House (n.d., index of political rights and civil liberties), Vanhanen (1997; index of democratization), Gupta, Jongman and Schmid (1994; observance of civil rights), Inglehart (1997, levels and stability of democracy). These indicators are all used to examine whether there is any systematic variation in response tendencies. All these control indicators are not assessed using rating scales and therefore can be used for our purposes, measuring socio-cultural and socio-political context variables at the nation-level (Table 1).

Finally, we used <u>indicators for prevalence rates</u> of various neuropsychiatric disorders as reported by Van Hemert et al. (2002b; derived from World Health Organization indicators). The disorders used were unipolar major depression, bipolar affective disorder, obsessive-compulsive disorders, panic disorders and psychosis. These variables are not based on ratings and capture psychological (and potentially biological) variability across cultural contexts.

Results

Prevalence of Acquiescence at Culture-Level

First, we correlated all indicators (including the previous indicators of acquiescent responding taken from the literature) to get an overall estimate of the shared variance. In line with other ecological analyses (e.g., Hofstede, 2001), we used Spearman rank-order correlation ($Rho \equiv \rho$) since this coefficient is not susceptible to outliers (which can have a big influence, particular in smaller samples). The expected value for the null hypothesis to be true would be .00. The observed mean overall correlation at the culture level was .096. This mean correlation is significantly higher than zero: $t(152) = 3.82$, 95% Confidence Interval: $.046 < \rho < .146$. This correlation indicates a small effect size according to Cohen (1988). Richard, Bond, & Stokes-Zoota (2003) reported that 30% of published studies in social psychology in the last 100 years reported a correlation of .10 or less, which have been interpreted as meaningful and substantive effects. The acquiescence indicator by McCrae et al. (2005) was derived from mean responses to balanced sets of positively and negatively phrased personality items. This may be a relatively pure indicator of *ARS* (Baumgartner & Steenkamp, 2001). The average correlation with all the variables on record is .081, which is very similar to the overall mean correlation in this study. It can be concluded that *CARS* can be seen as a consistent country characteristic that has a small, though relatively consistent influence on scores.

We also investigated patterns of correlations between previously reported indicators of acquiescence and our meta-analytic indicators. It is important to consider whether single response style indicators are actually pure indicators of response styles or whether they also capture some substantive variance (Baumgartner & Steenkamp, 2001; McCrae & Costa, 1983). An examination of these correlations suggests that previous indicators of acquiescence might have substantive meaning. For example, the source of guidance mean reported by Smith et al. (2002) was significantly correlated with state anxiety ($\rho = .44$) and the correlation with normative commitment also approached significance ($\rho = .41$, $p = .06$). This suggests that managers rely more on all sources of guidance in contexts where transitory anxiety is relatively high and people feel higher levels of pressure from social groups. Considering all sources of guidance might therefore help to alleviate this pressure and anxiety.

Second, the correlation with *the Schwartz Value Survey* (Schwartz, 2005) mean ratings also suggests some substantive meaning of the mean value ratings. Average value ratings at the country-level were significantly correlated with higher state anxiety ($\rho = .42$); greater identification with one's organization ($\rho = .43$), greater normative pressure from social groups ($\rho = .70$) and performance approach goals were marginally higher ($\rho = .49$, $p = .09$). This suggests a motivational component of the mean value rating. Commitment (both affective and normative) is often seen as an indicator of motivation. Together with the association with

performance approach goals, this pattern suggests that higher overall value ratings are indicative of greater motivation overall (which might be higher in contexts where there is a need to reduce anxiety). Fontaine, Duriez, Luyten, Corveleyn, and Hutsebaut (2005) reported similar results at the individual level. Therefore, our analysis suggests that single acquiescence indicators of positively phrased questions relating to specific concepts or domains (leadership guidance sources, values) could well have substantive meaning. It is therefore imperative to derive acquiescence indicators across a range of constructs and domains.

Explaining Acquiescence at Country Level

Having established that *CARS* is relatively stable, we then proceeded to explore the meaning of *CARS*. We computed *MARS,* the meta-analytically combined 'true' estimate of *CARS* across all reported indicators in Table 2. Collapsing across all indicators was supposed to minimize or even eliminate the influence of the specific domain carried by each indicator and therefore present a relatively pure estimate of *CARS*. One score was calculated for each country separately. Since there were missing data for various countries and indicators, we computed the mean if there were at least three indicators available for each of the countries. We arrived at an acquiescent response style estimate for 41 countries. Using Spearman rank-order correlations to avoid undue effect of extreme scores, a number of significant correlations between *MARS* and country-level indicators emerged (all reported effects are significant at $p < .05$ at least). *MARS* (as indicator of *CARS*) was related to survey quality derived from McCrae et al. (2005) ($\rho = -.50$), GDP ($\rho = -.39$), Purchasing Power Parity ($\rho = -.49$) and the Human Development Index ($\rho = -.55$). Income inequality was positively associated with *MARS* ($\rho = .38$). Population density and average annual GDP growth were not significantly correlated with *MARS* ($\rho < .30$).

The democracy indicators showed consistent findings across indicators. *MARS* was higher if citizens had less civil rights ($\rho = -.43$), civil liberties ($\rho = -.52$), and political rights ($\rho = -.52$) and *MARS* was lower if democratic institutions were more stable ($\rho = -.42$) and the level of democratization was higher (Vanhanen index: $\rho = -.63$; Humana index: $\rho = -.43$; level of democracy in 1990 and 1995: $\rho = -.46$; $\rho = -.53$, respectively).

Finally, a number of significant correlations with neuropsychiatric disorder prevalence rates emerged. *MARS* was associated with higher unipolar major depression rates ($\rho = .33$), bipolar affective disorder ($\rho = .39$), obsessive-compulsive disorder ($\rho = .40$) and panic disorder ($\rho = .42$). The correlation with psychosis prevalence rates was marginally significant ($\rho = .27$).

Discussion

These preliminary results suggest that the main effect of *CARS* is relatively small. The indicators used in this study came from a number of disciplines across a large number of independent samples and populations. The relatively low shared variance compared to some other studies that used secondary data analysis (e.g., Smith, 2004) is noteworthy. On one hand, it may suggest that means derived from published articles are less dependable (due to coding or typing errors, missing information in reported studies, etc.) and therefore our meta-analytic indicators may underestimate the prevalence of *ARS*. However, the number of significant and consistent correlations with other nation-level variables reported above suggests that these indicators have validity (see also Fischer & Chalmers, 2008, Fischer & Mansell, in press, Fischer & Smith, 2003; Dekker & Fischer, 2008; for the validity of meta-analytic approaches with means). *CARS* have a small, though consistent influence on the size of cross-cultural differences. Moreover, the patterning of our correlations suggests that differences between less and more affluent and less or more democratic countries are somewhat amplified by *CARS*.

Some of the previous *ARS* indicators have substantive meaning that increases correlations with other nation-level indicators. To derive independent indicators of response styles, data across a number of conceptually independent constructs would be needed. The current results suggest for example that both mean ratings across sources of guidance for

leadership decisions (Smith et al., 2002) and human values (Schwartz, 1994) might have some substantive meaning (in addition to capturing some acquiescence). Re-analyses of existing datasets focusing on specific topics (such societal or work-related issues; e.g., European Social Survey, Eurobarometer, organizational climate-type surveys) may overestimate the prevalence of *CARS* (or *IARS*).

Our present analysis just investigated the main effect of *CARS*. It is important that the interactions with method factors are also explored. This could be done by examining interactions with specific method factors (such as type of response scale or response labels used, number of response options, and clarity of items). This might even suggest a more substantial effect of *ARS* overall.

Exploring the Meaning of ARS at Culture Level

Our analysis indicates that the impact of *CARS,* although having only a relatively small main effect on survey responses, is rather systematic than random. This means that any shifts in means at the nation level have substantive meaning rather than being random error or bias. Table 1 outlined some potential correlates of *ARS* at individual and culture level. Our analysis included some of these indicators and they show an interesting and revealing picture of potential underlying processes of *CARS*. For example, increased survey means are associated with lower economic development, fewer experienced civil and political rights and lower levels of democracy. Similar correlations in magnitude and direction were also observed with the SVS, sources of guidance and acquiescence indicators derived from McCrae's et al. (2005) personality research (available from first author). Answering surveys has to be understood within the particular socio-cultural and socio-political context. At the nation level higher means to survey responses could mean that the socioeconomic conditions are lower and that democratic institutions are not very strong. Higher responses could be seen as a form of compliance. We could speculate that survey responses are systematic indicators of culturally appropriate expressiveness that are also reflected in the larger socio-political climate of the societies.

Higher mean levels to surveys are also associated with greater rates of displayed (and diagnosed) rates of psychiatric disorders. This is an intriguing finding which is stable across the various acquiescence indicators used here. Unipolar major depression, bipolar affective, obsessive-compulsive and panic order disorder rates are all positively and significantly correlated with the four different acquiescence indicators (2 correlations are marginally significant and 3 correlations out of 16 are not significant, but in the same direction, overall mean correlation = .31). A visual inspection of the correlations also suggests quite substantial levels of heteroscedasticity. Levels of *ARS* are around the midpoint of the scale for countries with lowest levels of prevalence rates. Increasing prevalence rates are then associated with higher *ARS* means but also greater variability of *ARS*. In nations with low levels of neuropsychiatric disorders, *ARS* is around the mean of the response scale (around .5 since all means were standardized to range from 0 to 1). With increasing levels of psychiatric disorders, the *ARS* mean increases, as does the spread of means, with some countries with high psychiatric disorder rates having very low *ARS* and others with similarly high levels of disorder rates having very high *ARS* rates. Two potential explanations could be put forward at this stage. First, these findings may reflect the economic development levels in such contexts (e.g., the extent to which countries have sufficient wealth to afford reliable administration systems and also allowing individuals access to health care so that they can be reliably diagnosed). This is a partial explanation since controlling for Human Development in 1987 leads to a reduction of observed correlations. The mean correlation drops from .31 to .21. Nevertheless, some of the correlations still remain significant (e.g., correlations between value-based acquiescence and unipolar/bipolar depression are least affected by partialling human development). Some correlations with *ARS* even increase, for example, the correlation between our meta-analytic

acquiescence index and rates of psychotic disorders increases and becomes highly significant, $r_{partial} = .99$, $p < .001$ (based on 31 countries). These findings therefore call into question the validity of these disorder rates. Since they are based on clinical examinations and judgments of experts within that particular cultural context, they could be an indicator of cultural processes (e.g., is it socially acceptable to diagnose individuals with disorders that may be seen as socially desirable or undesirable in a given cultural context?). It may point towards some culturally specific ways how partly biologically determined disorders are socially displayed and accepted. Our findings are preliminary and based on only one indicator of prevalence rates, however, the consistency of these results across a number of different variables, populations and methods point out that this area is awaiting much needed further exploration.

It is clear from these results, that survey responses from individuals coming from different cultural backgrounds may not be directly comparable, due to systematic variations in the environmental context of the individual. Researchers and practitioners need to consider this in their use of surveys across cultural groups. Even a small mean shift may have considerable implications for any decisions when the answers are evaluated without consideration of the particular context (e.g., when making decisions on policy issues that involves individuals from different cultural groups, decisions in court). However, it should also be noted that the current results only apply to results at the group level (culture level). No information about effects at the individual level can be made based on our analysis.

References

Bachman, J. G., & O'Malley, P. M. (1984). Yea-saying, nay-saying and going to extremes: Black-White differences in response styles. *Public Opinion Quarterly, 48,* 491-509.

Bass, B.M. (1985). *Leadership and performance beyond expectations.* New York: Free Press.

Baumgartner, H., & Steenkamp, J. B. E. M. (2001). Response styles in marketing research: A cross-national investigation. *Journal of Marketing Research, 37*, 143-156.

Boldt, E. D. (1976). Acquiescence and conventionality in a communal society. *Journal of Cross-Cultural Psychology, 7,* 21-36.

Carr, L. G., & Krause, N. (1978). Social status, psychiatric symptomatology and response bias. *Journal of Health and Social Behavior*, *19*, 86-91.

Crampton, S. M., & Wagner, J. A. (1994). Percept-percept inflation in micro-organizational research: An investigation on prevalence and effect. *Journal of Applied Psychology, 79*, 67-76.

Cheung, G. W., & Rensvold, R. B. (2000). Assessing extreme and acquiescence response sets in cross-cultural research using structural equations modeling. *Journal of Cross-Cultural Psychology, 31*, 160-186.

Chun, K-T., Campbell, J.B., & Yoo, J.H. (1974). Extreme response style in cross-cultural research: A reminder. *Journal of Cross-Cultural Psychology, 5*, 465-479.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349-354.

Dekker, S. & Fischer, R. (2008). Cultural Differences in Academic Motivation Goals: A meta-analysis across thirteen societies. *Journal of Educational Research, 101*, 99-110.

Elliot, A. J., & Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology, 72,* 218-232.

Fischer, R. (2004). Standardization to Account for Cross-Cultural Response Bias: A Classification of Score Adjustment Procedures and Review of Research in JCCP. *Journal of Cross-Cultural Psychology, 35*, 263-282.Fischer, R. (2008). Organizational Justice and Reward Allocation. In P. B. Smith, M. Peterson & D. Thomas (eds), *Handbook of Cross-Cultural Management, (pp. 135-150).* Los Angeles: Sage.

Fischer, R. & Chalmers, A. (2008). Is optimism universal? A meta-analytical investigation of optimism levels across 22 nations. *Personality and Individual Differences, 45,* 378-382.

Fischer, R., & Mansell, A. (in press). Levels of organizational commitment across cultures: A meta-analysis. *Journal of International Business Studies*.

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

Fontaine, J. R. J., Duriez, B., Luyten, P., Corveleyn, J., & Hutsebaut, D. (2005). Consequences of a multidimensional approach to religion for the relationship between religiosity and value priorities. *International Journal for the Psychology of Religion*, *15*, 123-143.

Freedom House (n.d.). *Annual survey of freedom country scores 1972-1973 to 1998-1999*. Washington, D.C. Author. Retrieved September 27, 1999 from http://www.freedomhouse.org/rankings.pdf

Georgas, J., & Berry, J. W. (1995). An ecocultural taxonomy for cross-cultural psychology. *Cross-Cultural Research, 29*, 121-157.

Guilford, J.P. (1954). Psychometric methods. (2nd ed.). New York: McGraw-Hill.

Goldberg, D. P. (1972). *The detection of psychiatric illness by questionnaire: A technique for the identification and assessment of non-psychotic psychiatric illness*. NY: Oxford University Press.

Greenleaf, E. A. (1992). Improving rating scale measures by detecting and correcting bias components in some response styles. *Journal of Marketing Research, 24*, 176-188.

Gupta, D. K., Jongman, A. J., & Schmid, A. P. (1994, July). *Assessing country performance in the field of human rights.* Paper presented at the XIII ISA Congress, Bielefeld, Germany.

Hofstede, G. (1980). *Culture's consequences: International differences in work-related values.* Beverly Hills: Sage.

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations* (2nd ed.). Thousand Oaks, CA: Sage.

House, R. J., Hanges, P. J., Javidan, M., Dorfman, P., & Gupta, V. (Eds.) (2003). *GLOBE, cultures, leadership, and organizations: GLOBE study of 62 societies.* Newbury Park, CA: Sage.

Hui, C.H., & Triandis, H.C. (1985). Measurement in cross-cultural psychology: A review and comparison of strategies. *Journal of Cross-Cultural Psychology, 16,* 131-152.

Hui, C.H., & Triandis, H.C. (1989). Effects of culture and response format on extreme response style. *Journal of Cross-Cultural Psychology, 20*, 296-309.

Humana, C. (1986). *World human rights guide.* London: The Economist Publications.

Inglehart, R. (1997). *Modernization and post-modernization. Cultural, economic and political change in 43 countries.* Princeton, NJ: Princeton University Press.

Johnson, T., Kulesa, P., Cho, Y. I., & Shavitt, S. (2005). The relation between culture and response styles: Evidence from 19 countries. *Journal of Cross-Cultural Psychology, 36,* 2, 264-277.

Lipsey, M.W., & Wilson, D.B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

Marin, G., Gamba, R.J., Marin, B.V. (1992). Extreme response style and acquiescence among Hispanics. *Journal of Cross-Cultural Psychology, 23*, 498-509.

McCrae, R. R., & Costa, P. T. (1983). Social desirability scales: More substance than style. *Journal of Consulting and Clinical Psychology, 51,* 882-888.

McCrae, R. R., Terracciano, A., 78 members of the Personality Profiles of Cultures Project. (2005). Universal features of personality traits from the observer's perspective: Data from 50 cultures. *Journal of Personality and Social Psychology, 88,* 3, 547-561.

Middleton, M. J., & Midgley, C. (1997). Avoiding the demonstration of lack of ability: An unexplored aspect of goal theory. *Journal of Educational Psychology, 89,* 710-718.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology,46,* 598-609.

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), Measures of personality and social psychological attitudes (pp. 17-59). San Diego, CA: Academic Press.

Phillips, D. L., & Clancy, K. J. (1970). Response bias in field studies of mental illness. *American Sociological Review, 35,* 503-515.

Ray, J. J. (1983). Reviving the problem of acquiescent response bias. *Journal of Social Psychology, 121*, 81-96.

Richard, F. D., Bond, C. F., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology, 7,* 331-363.

Ross, C. E., & Mirowsky, J. (1984). Socially-desirable response and acquiescence in a cross-cultural survey of mental health. *Journal of Health and Social Behavior, 25,* 189-197.

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton, NJ: Princeton University Press.

Rosenberg, M. (1979). *Conceiving the self.* New York: Basic Books.

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

Schwartz, S. H. (2004). Mapping and interpreting cultural differences around the world. In H. Vinken, J. Soeters, & P. Ester (Eds.), *Comparing cultures, dimensions of culture in a comparative perspective* (pp. 43-73). Leiden, the Netherlands: Brill.

Schwartz, S. H. (2005). Basic human values: Their content and structure across countries. In A. Tamayo & J. B. Porto (Eds.), *Valores e comportamento nas organizações* [Values and behavior in organizations] pp. 21-55. Petrópolis, Brazil: Vozes.

Smith, P. B. (2004). Acquiescent response bias as an aspect of cultural communication style. *Journal of Cross-Cultural Psychology, 35,* 50-61.

Smith, P. B., Peterson, M. F., Schwartz, S. H., Ahmad, A.H., Akande, D., Andersen, J.A., et al. (2002). Cultural values, sources of guidance and their relevance to managerial behavior: A 47 nation study. *Journal of Cross-Cultural Psychology, 33,* 188-208.

Smith, P. B., & Fischer, R. (2008). Acquiescence, Extreme Response Bias and Culture: A Multilevel Analysis. In F. J. R. van de Vijver, D.A. van Hemert & Y. H. Poortinga (Eds.) *Individuals and cultures in multi-level analysis* (pp. 285-314). New York: Erlbaum.

Spielberger, C.D., Gorsuch, R.L., & Lushene, R. E. (1970). The state-trait anxiety inventory (test manual). Palo Alto, CA: Consulting Psychologists Press.

Spector, P. E., & Brannick, M. T. (1995). The nature and effects of method variants in organizational research. In C. L. Cooper & I. T. Robertson (eds), *Internacional Review of Occupational and Organizational Psychology* (Vol 10, pp. 249-274). New York: John Wiley.

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

Van Dijk, T. K., Datema, F., Piggen, A-L. J. H. F., Welten, S. C. M., & Van de Vijver, F. J. R. (2009). Toward a Model of Acquiescence and Extremity Scoring for Cross-Cultural Research. In A. Gari & K. Mylonas (Eds.) *Q.E.D.: From Herodotus' Ethongraphic Journeys to Cross-Cultural Research* (pp. 149-158). Athens: Pedio.

Van Hemert, D. A., Van de Vijver, F. J. R., Poortinga, Y. H., & Georgas, J. (2002). Structural and functional equivalence of the Eysenck Personality Questionnaire within and between countries. *Personality and Individual Differences, 33,* 1229-1249.

Van Hemert, D. A., Van de Vijver, F. J. R., & Poortinga, Y. H. (2002). The Beck Depression Inventory as a measure of subjective well-being: A cross-national study. *Journal of Happiness Studies, 3,* 257-286.

Van Herk, H., Poortinga, Y. H., & Verhallen, T. M. M. (2004). Response styles in rating scales: Evidence of method bias in data from six EU countries. *Journal of Cross-Cultural Psychology,* 35, 346-350.

Vanhanen, T. (1997). *Prospects of democracy: A study of 172 countries*. London: Routledge.

Winkler, J. D., Kanouse, D. E., & Ware, J. E. (1982). Controlling for acquiescence response set in scale development. *Journal of Applied Psychology,* 67, 555-561.

Yu, J., & Murphy, K. R. (1993). Modesty bias in self-ratings of performance: A test of the cultural relativity hypothesis. *Personnel Psychology, 46,* 357-363.