



Technical Brief for the

THOMAS-KILMANN CONFLICT MODE INSTRUMENT

Global Research Project for Translated TKI Assessments

Craig A. Johnson
Richard C. Thompson
Michael G. Anderson
CPP Research Department



800-624-1765 | www.cpp.com

CONTENTS

Introduction 1

The Thomas-Kilmann Conflict Mode Instrument 1

Scoring 1

U.S. Norm Sample 1

Steps in This Study 2

Translation Process 2

Sample Selection 2

Social Desirability of the TKI Scales by Language 3

Reliability of the TKI Scales by Language 4

Validity of the TKI Scales by Language 6

Percentile Ranks from the International Samples 6

Analyses of Conflict Mode Differences 7

Conclusion 10

References 12

Appendix A: TKI Raw Scores and Percentiles 13

Appendix B: Tukey's HSD Post Hoc Analyses 16

Appendix C: TKI Mode Mean Score Differences
Between Men and Women 19

INTRODUCTION

The *Thomas-Kilmann Conflict Mode Instrument* (TKI) identifies a person's behavior in conflict situations. Understanding one's conflict-handling style and the broader range of conflict modes can help individuals manage conflict. The TKI is commonly used in a variety of applications, including management and supervisory training, team building, leadership coaching, and marriage and family counseling. CPP, Inc., in response to increasing demand for the TKI from outside the United States, has steadily translated the instrument from English into other languages. This technical brief summarizes translation efforts and reviews psychometric functioning in each language.

THE THOMAS-KILMANN CONFLICT MODE INSTRUMENT

The TKI, developed by Kenneth W. Thomas and Ralph H. Kilmann, is based on a conceptual framework proposed by Robert Blake and Jane Mouton (1964). Blake and Mouton suggested five basic attitudes and styles of control for managers that could be placed on a managerial grid (Blake & Mouton, 1964). Thomas refined the conflict aspects of this framework in the early 1970s (Thomas, 2002). The revised TKI conflict model identifies five different conflict-handling modes, or ways of dealing with conflict: *competing*, *collaborating*, *compromising*, *avoiding*, and *accommodating* (Schaubhut, 2007). The TKI describes these modes along two independent dimensions: *assertiveness*, the degree to which individuals attempt to satisfy their own concerns, and *cooperativeness*, the degree to which individuals attempt to satisfy other people's concerns (Thomas & Kilmann, 1974, 2007). As described in *Introduction to Conflict Management* (Thomas, 2002), the conflict modes represent the following major combinations of assertiveness and cooperativeness:

- *Competing* is assertive and uncooperative; individuals who use this mode try to satisfy their own concerns at the other person's expense.
- *Collaborating* is assertive and cooperative; individuals who use this mode try to find a win-win solution that completely satisfies both persons' concerns.
- *Compromising* is intermediate in assertiveness and cooperativeness; individuals who use this mode try to find an acceptable settlement that only partially satisfies both persons' concerns.
- *Avoiding* is unassertive and uncooperative; individuals who use this mode sidestep the conflict without trying to satisfy either person's concerns.

- *Accommodating* is unassertive and cooperative; individuals who use this mode attempt to satisfy the other person's concerns at the expense of their own.

A research review by Thomas and Kilmann (1975) found that other measures of the five modes of dealing with conflict were susceptible to respondents choosing responses based on how they may appear to others (Edwards, 1957). Research into socially desirable responding suggests a decrease in test validity if it is not controlled for (Edwards, 1970). Thomas and Kilmann developed the TKI to minimize disparity of social desirability in the item responses. They chose and tested each item pair to ensure that neither response was more socially desirable. A study examining the ability of the instrument to control social desirability found that the TKI significantly reduced the social desirability response bias when compared to similar tools assessing conflict behavior (Kilmann & Thomas, 1977; for a copy of this paper, see "References"). By actively addressing social desirability, the TKI more accurately describes individuals' conflict modes. A review of current literature found no studies of social desirability pertaining to newer assessments of the five conflict-handling modes.

SCORING

The TKI comprises 30 forced-choice items. For each item, respondents choose which of two statements is most characteristic of their behavior. Each conflict mode is paired with the remaining conflict modes, three times. Raw scores are calculated by counting the number of times each mode is chosen, with scores ranging from 0 to 12. The raw scores are converted to percentile scores. Percentile scores indicate the percentage of people in a norm group who scored at or below a given raw score. In addition, percentile scores are partitioned into three interpretive ranges—high (the top 25%), medium (the middle 50%), and low (the bottom 25%). Further information on percentile scores creation appears in the section "Percentile Ranks from the International Samples."

U.S. NORM SAMPLE

In this document, researchers used the 2007 U.S. norm sample as a comparative sample. Full descriptions of this sample and other U.S. samples are available in Schaubhut, 2007. Briefly, the 2007 U.S. norm sample included 8,000 adults from the U.S. employed population (Schaubhut, 2007). Respondents represented a number of racial and ethnic groups and self-reported working in a variety of occupations. Individuals in the sample mirror the U.S. workforce as

TABLE 1. TARGET LANGUAGES AND SAMPLE COUNTRIES

Language	Sample Country
English, North American	United States
Chinese, Simplified	China—Mainland
Chinese, Traditional	China—Hong Kong
Danish	Denmark
Dutch	Netherlands
French, Canadian	Canada
French, European	France
German	Germany
Portuguese, Brazilian	Brazil
Portuguese, European	Portugal
Spanish, European	Spain
Spanish, Latin American	Mexico
Swedish	Sweden

reported by the Bureau of Labor Statistics (U.S. Department of Labor, n.d.).

STEPS IN THIS STUDY

This research effort evaluated the translations of the TKI from North American English into 12 target languages (see Table 1). Researchers developed a research plan closely following Thomas and Kilmann’s original methods, supplementing it to address concerns unique to translations. Outlined below are the nine steps used in the study.

1. A translation consulting firm and an in-country subject matter expert translated the TKI into each target language.
2. A single sample was drawn from each target country and then broken into four subsamples. One subsample was used to examine social desirability, a second for test-retest, a third for validation, and a fourth for creation of norms.
3. Data were gathered from the four subsamples.
4. Researchers evaluated social desirability for each language.
5. Researchers examined test-retest data for each language.
6. Researchers examined validity of the TKI with the *Myers-Briggs Type Indicator*® (MBTI®) assessment for each language. Although it represents only one indication of validity, the MBTI tool provides a particularly rich research opportunity, as similar validations have been done using U.S. samples (Kilmann & Thomas, 1975).
7. Percentile ranks were generated for each international sample.
8. Conflict mode differences by language were examined.
9. Conflict mode differences by gender were examined.

The remaining sections of this document outline these steps in greater detail.

TRANSLATION PROCESS

TKI translation followed the practices used for other CPP translations. Most translations evaluated were completed prior to the beginning of this project; only the German translation was new. In all cases a translation consulting firm completed an initial translation. Next, an in-country subject matter expert reviewed the translation. Identified discrepancies were reconciled through review iterations between the translation firm and the in-country subject matter expert. In-country subject matter experts were given tie-breaking authority based on local dialect expertise. Once discrepancies were reconciled, the translations were placed online and reviewed for typographical errors.

SAMPLE SELECTION

Researchers partnered with an international marketing firm to identify participants in the 12 target languages. The marketing firm used an internal database to identify potential participants. The marketing firm contacted identified individuals via email and asked them to participate. Interested individuals responded and were screened for eligibility. Individuals were considered eligible if they were 18 years or older, self-reported being a full-time employee, and responded that their country of residence was the same as the country in which the sample was drawn. The sampling plan included an equal number of men and women in each language sample.

Researchers drew a master sample for each language, and all individuals completed the TKI. They then separated the master sample into four subsamples, with each completing another component of the research plan. The first subsample consisted of approximately 100 individuals per language who completed demographic information and rated the social desirability of TKI items. A second subsample consisted of approximately 400 individuals per language. These individuals completed demographic information and the TKI. A third subsample consisted of approximately 100 individuals per language who completed the TKI and then completed the MBTI assessment. Finally, approximately 100 individuals per language completed the TKI a second time after two weeks.

Researchers cleaned data before analyses, removing respondents if they omitted three or more TKI items, completed the assessment in three or fewer minutes, or demonstrated patterns of invalidity (e.g., responding to the survey with all

1s, 2s, or 0s). They scored remaining data using standard procedures. Table 2 provides a description of the final master international samples used in analyses.

SOCIAL DESIRABILITY OF THE TKI SCALES BY LANGUAGE

Individuals often respond to items on a self-report psychological inventory in a way they believe others view favorably. This tendency, commonly known as *social desirability bias*, is a persistent challenge for psychological measurement. Most commonly, socially desirable responding distorts measurement results by artificially increasing mean self-report scores on desirable attributes (Thomas & Kilmann, 1975).

The TKI is a forced-choice measure, where the respondent selects one of two alternatives, making the social desirability of response options an important consideration. Thomas and Kilmann designed the TKI to account for social desirable responding by having participants rate the response alternatives of the TKI instrument on a 9-point scale, ranging from “Extremely Undesirable” to “Extremely Desirable” (see Edwards, 1957). Thomas and Kilmann paired response alternatives according to the rated social desirability.

Kilmann and Thomas (1977) also used social desirability ratings to determine the influence of item desirability on item and scale scores. First, they created a social desirability differential for each item by subtracting the mean social desirability rating of alternative “B” from alternative “A.” The authors then correlated the differential scores with the proportion of individuals who selected the “A” alternative when completing the TKI instrument. The results of this investigation indicated that social desirability did not have a significant impact on response alternative selection ($r = .21$; $p > .01$; $n = 30$), explaining only 4% of the variance in self-ratings on TKI instrument items.

Researchers replicated social desirability analyses in the current investigation for each translation of the TKI (see Table 3). Results for 9 of the 12 new translations indicated that social desirability showed no significant correlation with item response selection ($p > .05$). The findings show that socially desirable ratings explained less than 1% to, in the extreme case, 35% of the variance in self-ratings on TKI instrument items. It is important to note that even the most extreme case is well below the equivalent correlations for other conflict-handling instruments (.87 to .94; Kilmann & Thomas, 1977). Researchers also examined results to determine whether correlations significantly differed from the U.S. population using a Steiger’s *Z*-test. No statistically significant

TABLE 2. DESCRIPTION OF THE U.S. NORM SAMPLE AND INTERNATIONAL SAMPLES BY LANGUAGE

Language	Mean Age	Men %	Organizational Level				
			Entry Level %	Nonsupervisory %	Supervisory %	Management %	Executive %
English, North American	40.4	50.0	5.0	20.0	20.0	35.0	20.0*
Chinese, Simplified	32.6	48.1	21.6	25.9	30.9	18.9	2.7
Chinese, Traditional	33.2	48.1	45.2	29.7	16.3	5.3	3.4
Danish	45.5	48.6	41.3	15.9	25.4	14.9	2.5
Dutch	43.7	49.4	70.7	7.6	11.0	6.9	3.8
French, Canadian	43.1	49.3	59.0	12.5	11.5	10.8	6.2
French, European	38.7	48.5	52.3	11.2	25.2	5.0	6.2
German	40.0	48.2	68.6	12.9	7.3	4.6	6.6
Portuguese, Brazilian	36.9	47.2	61.1	15.4	15.4	6.8	1.3
Portuguese, European	37.6	50.7	46.3	11.8	21.2	15.3	5.5
Spanish, European	39.5	46.9	53.1	27.9	11.2	2.7	5.1
Spanish, Latin American	38.1	51.0	30.0	28.0	17.3	15.0	9.8
Swedish	47.3	48.2	60.3	26.1	9.1	2.1	2.4

Note: North American English $N = 8,000$, Simplified Chinese $N = 285$, Traditional Chinese $N = 318$, Danish $N = 350$, Dutch $N = 346$, Canadian French $N = 367$, European French $N = 324$, German $N = 330$, Brazilian Portuguese $N = 321$, European Portuguese $N = 369$, European Spanish $N = 339$, Latin American Spanish $N = 351$, Swedish $N = 355$.

*Percentage includes executives (15%) and top executives (5%) for this sample.

TABLE 3. CORRELATION BETWEEN SOCIAL DESIRABILITY DIFFERENTIAL AND PERCENTAGE OF RESPONSE OPTION ENDORSEMENT OF THE U.S. NORM SAMPLE AND INTERNATIONAL SAMPLES BY LANGUAGE

Language	<i>r</i>	U.S. Norm Sample Comparison		
		<i>Z</i>	Diff	<i>p</i>
English, North American ¹	.210	0.213	n/a	n/a
Chinese, Simplified	.592*	0.681	1.72	.086
Chinese, Traditional	.084	0.084	-0.47	.636
Danish	.111	0.111	-0.37	.709
Dutch	.436*	0.467	0.93	.350
French, Canadian	.099	0.099	-0.42	.676
French, European	.458*	0.495	1.03	.301
German	-.012	-0.012	-0.83	.408
Portuguese, Brazilian	.274	0.281	0.25	.803
Portuguese, European	.135	0.136	-0.28	.776
Spanish, European	-.019	-0.019	-0.85	.394
Spanish, Latin American	.300	0.310	0.35	.723
Swedish	-.034	-0.034	-0.91	.364

Note: *r* = correlation between social desirability differential and percentage of response option endorsement; *Z* = Steiger's *Z*, a measure of statistical difference between correlations; Diff = difference between the U.S. and comparison country, where positive values indicate scores higher in the comparison country; *p* = significance test, where values less than .05 are significant.

**p* < .05. ¹From Kilmann & Thomas, 1977.

differences existed between the U.S. norm sample and translated versions.

A second approach Kilmann and Thomas (1977) implemented involved averaging social desirability differentials between paired response alternatives for the 12 statements representing each of the five conflict-handling modes. These values were then correlated with the average self-report scores across the five modes, yielding a Pearson coefficient of .41 (nonsignificant for *n* = 5). Researchers also replicated these analyses with results reported in Table 4. Consistent with previous findings, 11 out of 12 samples indicate that social desirable responding did not have a significant influence on scores. European Spanish reached statistical significance (*p* = .02), indicating that social desirable responding may have had an influence on conflict-handling scores. The European Spanish sample was also the only sample that differed significantly from the U.S. norm sample.

RELIABILITY OF THE TKI SCALES BY LANGUAGE

Reliability of an assessment is commonly measured in two ways. First, items composing a scale are examined to deter-

mine item consistency using Cronbach's alpha. Higher values on Cronbach's alpha indicate greater consistency. As Thomas, Thomas, and Schaubhut (2008) point out, because only four degrees of freedom exist among five scale scores, Cronbach's alpha undervalues the reliability of the TKI scales. This fact is further complicated as responses for each item load onto different scales. Due to these factors and others, Thomas, Thomas, and Schaubhut conclude that Cronbach's alpha may not be the best indicator of reliability for the TKI. For these reasons researchers did not examine internal consistency using Cronbach's alpha. For further information on the discussion of internal consistency and the TKI, see Thomas, Thomas, and Schaubhut (2008).

A second evaluation of reliability looks at consistency of responses over time. Test-retest asks participants to complete a measure on multiple occasions and then correlates the scores. In the current analysis, participants completed the TKI and, after two weeks, completed the TKI again. Test-retest reliabilities for each translation are reported in Table 5. In general, the test-retest reliabilities for the translated versions are below those reported for the North American English version (Kilmann & Thomas, 1977), with Simplified Chinese and Traditional Chinese showing the lowest test-retest coefficients. Considering test-retest reliabilities across all languages, however, yields test-

TABLE 4. CORRELATION BETWEEN SOCIAL DESIRABILITY DIFFERENTIAL AND MODE SCORE FOR THE U.S. NORM SAMPLE AND INTERNATIONAL SAMPLES BY LANGUAGE

Language	<i>r</i>	U.S. Norm Sample Comparison		
		<i>Z</i>	Diff	<i>p</i>
English, North American ¹	.410	0.436	n/a	n/a
Chinese, Simplified	.600	0.693	0.26	.797
Chinese, Traditional	.307	0.317	-0.12	.906
Danish	-.772	-1.025	-1.46	.144
Dutch	-.011	-0.011	-0.45	.655
French, Canadian	.166	0.168	-0.27	.789
French, European	-.291	-0.300	-0.74	.462
German	-.322	-0.334	-0.77	.442
Portuguese, Brazilian	-.277	-0.284	-0.72	.471
Portuguese, European	-.137	-0.138	-0.57	.566
Spanish, European	-.937*	-1.713	-2.15	.032
Spanish, Latin American	-.213	-0.216	-0.65	.514
Swedish	-.831	-1.191	-1.63	.104

Note: *r* = correlation between social desirability differential and mode score; *Z* = Steiger's *Z*, a measure of statistical difference between correlations; Diff = difference between the U.S. and comparison country, where positive values indicate scores higher in the comparison country; *p* = significance test, where values less than .05 are significant.

**p* < .05. ¹From Kilmann & Thomas, 1977.

TABLE 5. TEST-RETEST CORRELATIONS FOR THE U.S. NORM SAMPLE AND INTERNATIONAL SAMPLES BY LANGUAGE

Language	<i>N</i>	Competing	Collaborating	Compromising	Avoiding	Accommodating
English, North American ¹	76	.61	.63	.66	.68	.62
Chinese, Simplified	78	.38	.15	.35	.30	.40
Chinese, Traditional	91	.62	.44	.28	.25	.47
Danish	95	.71	.57	.45	.55	.51
Dutch	99	.72	.57	.46	.58	.55
French, Canadian	59	.70	.46	.43	.56	.57
French, European	64	.42	.33	.46	.40	.50
German	77	.50	.63	.34	.51	.52
Portuguese, Brazilian	92	.57	.23	.62	.47	.40
Portuguese, European	80	.54	.49	.66	.57	.57
Spanish, European	85	.63	.46	.43	.33	.52
Spanish, Latin American	78	.59	.46	.53	.43	.63
Swedish	113	.81	.51	.53	.49	.57

Note: See Kilmann & Thomas, 1977, for North American English correlations. All remaining correlations greater than .15 are significant at the .05 level.

¹From Kilmann & Thomas, 1977.

retest coefficients comparable with those of other conflict mode assessments (.39 to .55; Kilmann & Thomas, 1977).

VALIDITY OF THE TKI SCALES BY LANGUAGE

Validation of an instrument establishes empirical links to theoretical constructs and established measures. Kilmann and Thomas (1975, 1977) explored the empirical links between the TKI and several other assessments. In this study, researchers sought to replicate known and meaningful relationships between the TKI and the preference dichotomies of the MBTI instrument (Kilmann & Thomas, 1975). Kilmann and Thomas (1975) found that the *Integrative* dimension of conflict behavior, which refers to the realization of both parties' interest, was positively correlated with the Extraversion–Introversion dichotomy ($r = .29, p < .01$), indicating that Introverts were more likely to use Avoiding than they were Collaborating conflict behaviors. The *Distributive* dimension, which refers to attempts to satisfy one's own objectives at the expense of the other party, was negatively related to the Thinking–Feeling dichotomy ($r = -.38, p < .001$), indicating that individuals with a Feeling preference were more likely to use Accommodating than they were Competing conflict behaviors.

In this study, a subsample of participants completed the TKI followed by the MBTI assessment. The Integrative dimension, consistent with Kilmann and Thomas (1975), subtracted the Avoiding score from the Collaborating score. The Distributive dimension subtracted the Accommodating score from the Competing score. Building on previous research, the current investigation explored the following hypotheses:

1. The Integrative dimension will be correlated with the Extraversion–Introversion dichotomy.
2. The Distributive dimension will be correlated with the Thinking–Feeling dichotomy.
3. Competing conflict behaviors, because they involve satisfying personal needs at the expense of others, will be correlated with the Thinking–Feeling dichotomy.
4. Collaborating conflict behaviors, because they require a deeper level of involvement with the other party, will be correlated with the Extraversion–Introversion dichotomy.
5. Avoiding conflict behaviors, because they require the least amount of interpersonal interaction, will be correlated with the Extraversion–Introversion dichotomy.

6. Accommodating conflict behaviors, because they involve concern for the needs of others, will be correlated with the Thinking–Feeling dichotomy.

Hypothesized relationships were investigated using Pearson product-moment correlation coefficients and are shown in Table 6. Statistically significant results supported the hypotheses for the Danish, Canadian French, European French, and Swedish samples. Dutch, German, Brazilian Portuguese, European Spanish, and Latin American Spanish samples generally showed the expected patterns of relationships but did not reach statistical significance for all relationships. Simplified Chinese and European Portuguese samples showed smaller relationships, with trends in the expected direction except for hypothesis 4. Traditional Chinese results were less consistent, with results in the opposite direction of hypotheses 4 and 5.

PERCENTILE RANKS FROM THE INTERNATIONAL SAMPLES

Researchers calculated percentiles for each of the raw scores (0–12). Percentile scores are calculated as the median point (or middle) of the range of cumulative frequency covered by that score. For example, if a raw score of 5 has a cumulative frequency of 40% and a score of 6 has a cumulative frequency of 60%, then a raw score of 6 covers the range from 40% to 60% and the percentile assigned would be the median value of 50%. If researchers did not make this adjustment and used simple cumulative frequencies, the percentile scores would be biased in an upward direction. Among other things, this would mean that more than 25% of people would be expected to fall into the “top 25%” interpretive range on a given conflict mode, while fewer than 25% of people would be expected to fall into the “bottom 25%” interpretive range.

The raw and percentile scores for the U.S. norm sample and each of the translated versions of the TKI appear in Tables A-1 to A-5 in Appendix A. As indicated in these tables, the percentiles tend to shift slightly across translated versions of the assessment. To interpret these shifts, it is important to understand that a percentile score denotes the number of individuals below the raw score. For example, a percentile score of 70% indicates that 70% of the normative group scores at or below that raw score.* A review of the percentile scores shows a relatively consistent pattern of percentiles across countries. A noteworthy exception is a shift on Collaborating in comparison to the U.S. sample. Percentiles on this scale indicate

*A difference in interpretation of values exists when comparing mean or median scores on a mode between two or more different samples and when comparing the same samples' percentile values for the same raw score (as in Table 7, as well as in Tables A-1 to A-5). When comparing the median or mean raw scores on, for example, Competing between two samples—sample A and sample B—a higher score for sample A means that sample A scored higher on the raw scores for Competing. But when comparing the percentile values of the two samples for the same raw scores, a higher percentile value means the opposite. For example, when sample A has a higher percentile than sample B for the same raw score, it means that more people in sample A have scored lower than that raw score.

TABLE 6. CORRELATIONS BETWEEN TKI MODES AND MBTI® CONTINUOUS SCORES FOR THE U.S. NORM SAMPLE AND INTERNATIONAL SAMPLES BY LANGUAGE

Language	N	TKI Dimensions		TKI Modes			
		Integrative	Distributive	Competing	Collaborating	Avoiding	Accommodating
		MBTI® Dichotomies		MBTI® Dichotomies			
		E-I	T-F	T-F	E-I	E-I	T-F
English, North American ¹	76	-.29**	-.38**	n/a	n/a	n/a	n/a
Chinese, Simplified	62	-.10	-.10	-.08	.01	.17	.09
Chinese, Traditional	73	.12	-.38**	-.27**	.01	-.18	.35**
Danish	93	-.28**	-.29**	-.31**	-.27**	.21*	.18*
Dutch	11	-.20*	-.27**	-.27**	-.11	.22*	.21*
French, Canadian	43	-.41**	-.39**	-.39**	-.34*	.36**	.27*
French, European	69	-.35**	-.33**	-.30**	-.24*	.37**	.27*
German	60	-.21	-.32**	-.29*	-.20	.16	.25*
Portuguese, Brazilian	77	-.15	-.03	-.03	-.20*	.05	.02
Portuguese, European	86	-.01	-.10	-.09	.01	.02	.08
Spanish, European	10	-.16*	-.11	-.08	-.06	.21*	.12
Spanish, Latin American	85	-.12	-.22*	-.24*	-.06	.14	.12
Swedish	11	-.27**	-.39**	-.45**	-.18*	.26**	.20*

Note: Compromising is not included for analysis, as it is in the middle of the TKI modes.

* $p < .05$; ** $p < .01$. ¹From Kilmann & Thomas 1977.

that individuals in other countries tended to have lower raw scores.

Also in Appendix A, Table A-6 contains the conflict mode percentile medians by language using U.S. norms. As shown in this table, medians tend to be near the center of the distribution, indicating that half of the distribution is above the median and half below the median.

As with the U.S. norm sample, each international sample was partitioned into three interpretive categories—high (the top 25%), medium (the middle 50%), and low (the bottom 25%), based on the distribution of the raw scores. Table 7 illustrates the results for each of the five conflict modes. Although each language sample’s percentile scores shifted slightly for all conflict modes, the interpretive categories are very similar to those found for the U.S. norm sample, with translations normally varying by one or two raw points. Thus, interpretation of TKI results for the U.S. norm sample and the international samples are similar. The one possible exception is the Collaborating mode, where ranges of 5–8 and 9–12 were categorized as medium and high, respectively, for the U.S. norm sample and ranges of 5–7 and 8–12 were

categorized as low and medium for some other countries. This difference may be a function of the organizational level of individuals in the U.S. norm sample versus that of those in the international samples. In the U.S. norm sample 75% of respondents were employed as supervisors, managers, executives, or top executives. In contrast, in the international samples on average only about 30% of respondents were employed as supervisors, managers, executives, or top executives. Previous research (Blake & Mouton, 1964; Brewer, Mitchell, & Weber, 2002; Thomas, Thomas, & Schaubhut, 2008) has indicated that Collaborating increases at higher organizational levels. Given these findings, the shift in the Collaborating mode in the current study may be attributed to the smaller percentage of individuals in the international samples who were supervisors or above.

ANALYSES OF CONFLICT MODE DIFFERENCES

The analyses of the raw score and percentile distributions show slight differences in the distributions and interpretive boundaries for several languages. However, it is necessary to

**TABLE 7. RAW SCORES AND INTERPRETIVE RANGES FOR THE U.S. NORM
SAMPLE AND INTERNATIONAL SAMPLES BY LANGUAGE**

TKI Mode	English, North American	Chinese, Simplified	Chinese, Traditional	Danish	Dutch	French, Canadian	French, European	German	Portuguese, Brazilian	Portuguese, European	Spanish, European	Spanish, Latin American	Swedish
Competing													
High (top 25%)	7–12	8–12	7–12	7–12	8–12	7–12	8–12	6–12	7–12	7–12	7–12	7–12	7–12
Medium (middle 50%)	3–6	4–7	3–6	3–6	3–7	3–6	3–7	2–5	3–6	3–6	3–6	4–6	2–6
Low (bottom 25%)	0–2	0–3	0–2	0–2	0–2	0–2	0–2	0–1	0–2	0–2	0–2	0–3	0–1
Collaborating													
High (top 25%)	9–12	8–12	7–12	8–12	7–12	8–12	8–12	8–12	7–12	8–12	7–12	7–12	7–12
Medium (middle 50%)	5–8	5–7	5–6	5–7	4–6	5–7	5–7	5–7	5–6	5–7	4–6	4–6	4–6
Low (bottom 25%)	0–4	0–4	0–4	0–4	0–3	0–4	0–4	0–4	0–4	0–4	0–3	0–3	0–3
Compromising													
High (top 25%)	10–12	9–12	9–12	8–12	9–12	10–12	10–12	10–12	10–12	10–12	9–12	10–12	9–12
Medium (middle 50%)	6–9	6–8	6–8	5–7	5–8	7–9	7–9	7–9	7–9	7–9	6–8	7–9	6–8
Low (bottom 25%)	0–5	0–5	0–5	0–4	0–4	0–6	0–6	0–6	0–6	0–6	0–5	0–6	0–5
Avoiding													
High (top 25%)	8–12	7–12	8–12	9–12	10–12	8–12	7–12	8–12	8–12	9–12	9–12	9–12	10–12
Medium (middle 50%)	5–7	5–6	5–7	6–8	6–9	5–7	5–6	5–7	5–7	5–8	6–8	6–8	6–9
Low (bottom 25%)	0–4	0–4	0–4	0–5	0–5	0–4	0–4	0–4	0–4	0–4	0–5	0–5	0–5
Accommodating													
High (top 25%)	7–12	8–12	9–12	8–12	9–12	8–12	8–12	8–12	7–12	7–12	9–12	8–12	9–12
Medium (middle 50%)	4–6	5–7	5–8	5–7	5–8	4–7	4–7	5–7	4–6	3–6	5–8	4–7	5–8
Low (bottom 25%)	0–3	0–4	0–4	0–4	0–4	0–3	0–3	0–4	0–3	0–2	0–4	0–3	0–4

Note: Interpretive ranges that differ between the U.S. norm sample and the international samples are shaded.

determine whether these differences are meaningful to users of the TKI. To that end, a series of analyses examined differences in mean scores based on language and gender.

Differences by Language

Researchers used univariate analyses of variance (ANOVAs) to determine whether statistically significant differences existed between samples from the U.S. and the other 12 languages. ANOVAs compare the mean scores of two or more groups to determine whether statistically significant differences exist between them (Tabachnick & Fidell, 2001). A summary of these analyses provided in Table 8 shows that statistically significant mean differences exist between languages for all five conflict modes. However, given the sample sizes, the more relevant criteria for this analysis are the effect sizes (η^2).

Researchers computed effect sizes to determine the magnitude of the ANOVA results. Effect sizes provide an indication of size differences in a practical sense. While results can be statistically significant due to large sample sizes, they may not have practical significance. Large samples typically reach statistical significance, so effect sizes are of particular value in this study. Effect sizes computed for the mean differences on the five TKI conflict modes suggest that the reported differences are small and not likely to have a large impact on any of the five conflict modes (Cohen, 1992). The effect sizes for the overall comparison are also shown in Table 8. A review shows that the largest effect size for the comparison of all countries occurs for Collaborating, accounting for just under 6% of the variance in Collaborating across all the different languages examined. An η^2 of this value is a small effect, according to Cohen (1992).

After the ANOVAs showed statistically significant differences between the countries, Tukey's HSD post hoc analyses evaluated test differences between the U.S. norm sample and each language sample. Standardized difference scores, reported in Appendix B, examined the magnitude of these differences. Effect sizes measured using Cohen's d revealed that differences were small ($d \leq .20$) to medium ($d \geq .50$ and $< .80$) (Cohen, 1992). Notable exceptions were mean comparisons between the U.S. norm sample and the Dutch, Brazilian Portuguese, European Spanish, Latin American Spanish, and Swedish translations on the Collaborating scale, which demonstrated medium ($d \geq .50$) to high ($d \geq .80$) effect sizes. These differences may exist due to a different percentage of supervisors and executives in the U.S. norm sample in comparison to each international sample.

Differences by Gender

Previous research has revealed differences between the scores of men and women in the U.S., with men scoring slightly higher than women on Competing (Thomas, Thomas, & Schaubhut, 2008). Herk, Thompson, Thomas, and Kilmann (2011) found that international samples followed a similar pattern when completing the North American English version of the TKI, with 10 out of 17 countries having men scoring higher on Competing. Gender differences on other modes were less evident. In this research project, analyses were replicated to determine whether similar differences existed. Table 9 provides median percentile scores for men and women using U.S. percentiles in each of the countries for each of the five conflict modes.

To examine gender differences by language, researchers used exploratory independent sample t -tests (see Appendix C).

TABLE 8. ANOVA RESULTS FOR THE TKI MODES FROM THE U.S. NORM SAMPLE AND INTERNATIONAL SAMPLES

TKI Mode	Variance Sources	Sum of Squares (<i>SS</i>)	Degrees of Freedom (<i>df</i>)	Mean Square (<i>MS</i>)	<i>F</i>	<i>p</i>	η^2
Competing	Language	537.467	12	44.789	5.695	.000	.0056
	Error	95242.844	12042	7.865			
Collaborating	Language	3392.595	12	282.716	63.627	.000	.0596
	Error	56899.335	12042	4.443			
Compromising	Language	1298.467	12	108.206	23.258	.000	.0227
	Error	57322.669	12042	4.652			
Avoiding	Language	1660.107	12	138.342	25.604	.000	.0249
	Error	66725.373	12042	5.403			
Accommodating	Language	1473.454	12	122.788	24.312	.000	.0237
	Error	62291.086	12042	5.050			

TABLE 9. TKI MODE MEDIAN PERCENTILE SCORES FOR MEN AND WOMEN IN INTERNATIONAL SAMPLES BY LANGUAGE USING U.S. NORMS

Language	Competing		Collaborating		Compromising		Avoiding		Accommodating	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
	Median Percentile									
Chinese, Simplified	57	57	41	41	41	41	34	49	62	62
Chinese, Traditional	57	44	41	41	41	58	49	49	62	76
Danish	44	44	41	41	27	27	65	65	62	62
Dutch	57	44	26	26	41	41	65	65	62	62
French, Canadian	57	44	41	41	58	58	34	49	46	46
French, European	57	57	41	41	58	58	34	49	46	62
German	44	31	41	41	58	58	49	65	62	62
Portuguese, Brazilian	57	57	26	26	58	58	49	49	62	62
Portuguese, European	57	44	41	41	41	58	49	49	46	46
Spanish, European	44	44	26	26	41	41	65	65	62	76
Spanish, Latin American	57	57	26	26	58	58	65	65	46	46
Swedish	57	31	26	26	41	41	65	65	76	76

Independent sample *t*-tests examine whether significant differences exist between two populations (Gravetter & Wallanu, 2004). Researchers also used Cohen’s *d* to examine effect sizes to determine practical significance. Results showed that for Competing, men scored significantly higher in the Danish, European Spanish, and Swedish samples. Men also scored significantly higher on the Collaborating mode in the European Portuguese and Latin American Spanish samples. German women scored higher than men on the Compromising mode. Latin American Spanish women scored higher on the Avoiding mode. All effect sizes were less than or equal to .26, indicating small practical differences.

In general, few differences were found between men’s and women’s mode scores. Differences identified were small in magnitude and were generally consistent with previous research (Herk et al., 2011; Thomas, Thomas, & Schaubhut, 2008).

CONCLUSION

The analyses presented in this document were conducted as an initial investigation of TKI translations in multiple languages. Researchers first evaluated social desirability. To begin, they evaluated the correlations between social desirability and percentage of response option endorsement. Results showed social desirability responding in Simplified Chinese, Dutch, and European French samples, with the

remaining nine samples showing no significant relationships. The relationship between social desirability differential and percentage of response option endorsement was then compared with results seen in the U.S. sample, with no significant differences found. Researchers also evaluated the correlation between social desirability and mode score. Results showed no significant differences for 11 out of the 12 language samples, with European Spanish being the sole exception. These results were again compared to the U.S. sample. Results showed similar patterns, with 11 out of 12 of these correlations not being significantly different from relationships seen in the U.S. sample. The sole exception found was the European Spanish sample. As a whole, these results indicate that social desirability is not a large concern for the evaluated samples. The influence found does not appear to significantly impact scale scores and is much lower than that reported for other conflict mode instruments. Furthermore, there appears to be a consistent pattern of social desirability across languages.

Researchers also evaluated the test-retest reliabilities of the TKI translations. Results showed values lower than those seen in the U.S. norm sample and below what are considered ideal for psychological measurement. However, test-retest reliabilities tend to be within the expected range for other conflict mode tools (.39 to .55; Kilmann & Thomas, 1977). Future research should evaluate whether these reliabilities are symptomatic of the samples, the research method, or problems with the translations.

Researchers examined the TKI norms, including the high, medium, and low categories. Results revealed no large deviations from the U.S. norms. However, broad generalizations about languages, countries, or cultures should not be made based on these findings.

A review of the external validity of the assessment in relation to the MBTI assessment was favorable. Several translations exhibited significant results in the expected direction; the remaining translations showed results in the expected direction but may not have reached statistical significance on all comparisons. Only the Traditional Chinese sample showed three patterns that were unexpected. Taken as a whole, these results provide evidence supporting the validity of the TKI conflict modes.

Researchers compared the mean conflict mode across languages. Analyses showed that most differences were small to medium for the Competing, Compromising, Avoiding, and Accommodating modes. Collaborating showed larger differences for Dutch, European Portuguese, European Spanish, Latin American Spanish, and Swedish samples.

Finally, researchers examined differences between men and women for each conflict mode. Results showed very few differences between men and women across the various languages, and these differences were found to be small to medium in size.

In general, based on social desirability, test-retest, and validity data, the translations of Danish, Dutch, Canadian French,

European French, German, Latin American Spanish, and Swedish closely mirror U.S. findings. European Portuguese results were consistent with social desirability and test-retest reliability with slightly weaker validity results with the MBTI assessment. These differences are likely due to sample size issues and should be replicated in a larger sample. Research also indicated that Simplified Chinese and Traditional Chinese translations more frequently deviated from the U.S. sample than did other languages, particularly with test-retest and validity-related issues. Future studies should look further into these languages. Brazilian Portuguese also showed some differences from the U.S. norm sample, with Collaborating test-retest and validity results lower but in the anticipated directions. Similarly, European Spanish differed slightly from the U.S. norm sample in regard to social desirability differential and mode score and Avoiding test-retest, and minor differences in validity evidence. It is important to note that the validity sample for this group was particularly small ($n = 10$), and trends were in the anticipated direction.

In general, results of these studies are supportive of the TKI translations. Results are generally consistent and of the same magnitude across countries. Future research should focus on two areas: first, research in the document should be replicated with larger sample sizes as they become available; and second, research should be conducted on translated versions of the TKI as they are developed and used around the world to ensure that the instrument functions well in those languages and cultures.

REFERENCES

- Blake, R., & Mouton, J. S. (1964). *The managerial grid*. Houston, TX: Gulf.
- Brewer, N., Mitchell, P., and Weber, N. (2002). Gender role, organizational status, and conflict management styles. *The International Journal of Conflict Management*, *13*, 78–94.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.
- Edwards, A. L. (1957). *The social desirability variable in personality assessment and research*. New York: Dryden, 1957.
- Edwards, A. L. (1970). *The measurement of personality traits by scales and inventories*. New York: Holt, Rinehart & Winston.
- Gravetter, F., & Wallnau, L. (2004). *Statistics for the behavioral sciences* (6th ed.). Belmont, CA: Wadsworth.
- Herk, N. A., Thompson, R. C., Thomas, K. W., & Kilmann, R. H. (2011). *International technical brief for the Thomas-Kilmann Conflict Mode Instrument*. Mountain View, CA: CPP, Inc. Available for download at https://www.cpp.com/Pdfs/TKI_Intl_Tech_Brief.pdf.
- Kilmann, R. H., & Thomas, K. W. (1977). Developing a forced-choice measure of conflict-handling behavior: The MODE instrument. *Educational and Psychological Measurement*, *37*(2), 309–325. Available for download at www.kilmanniagnostics.com/developing-forced-choice-measure-conflict-handling-behavior-mode-instrument.
- Kilmann, R. H., & Thomas, K. W. (1975). Interpersonal conflict-handling behavior as reflections of Jungian personality dimensions. *Psychological Reports*, *3*, 971–980.
- Schaubhut, N. (2007). *Technical brief for the Thomas-Kilmann Conflict Mode Instrument: Description of the updated normative sample and implications for use*. Mountain View, CA: CPP, Inc. Available for download at https://www.cpp.com/Pdfs/TKI_Technical_Brief.pdf.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. Needham Heights, MA: Allyn & Bacon.
- Thomas, K. W. (2002). *Introduction to conflict management: Improving performance using the TKI*. Mountain View, CA: CPP, Inc.
- Thomas, K. W., & Kilmann, R. H. (1974, 2007). *Thomas-Kilmann Conflict Mode Instrument*. Mountain View, CA: CPP, Inc.
- Thomas, K. W., & Kilmann, R. H. (1975). The social desirability variable in organizational research: An alternative explanation for reported findings. *Academy of Management Journal*, *18*(4), 741–752.
- Thomas, K. W., Thomas, G. F., & Schaubhut, N. (2008). Conflict styles of men and women at six organization levels. *The International Journal of Conflict Management*, *19*, 148–166.
- U.S. Department of Labor, Bureau of Labor Statistics (n.d.). Household data annual averages, table 18. Retrieved August 15, 2006; www.bls.gov/cps/cpsaat18.pdf.

APPENDIX A: TKI RAW SCORES AND PERCENTILES

TABLE A-1. RAW SCORES AND PERCENTILES FOR THE U.S. NORM SAMPLE AND INTERNATIONAL SAMPLES—COMPETING

Raw Score	Percentile												
	English, North American	Chinese, Simplified	Chinese, Traditional	Danish	Dutch	French, Canadian	French, European	German	Portuguese, Brazilian	Portuguese, European	Spanish, European	Spanish, Latin American	Swedish
0	3	1	2	4	4	2	1	4	2	2	5	1	4
1	10	3	8	13	12	7	7	15	7	7	14	6	14
2	20	10	18	24	23	16	16	27	17	14	24	14	26
3	31	19	29	34	35	28	27	41	28	26	34	21	39
4	44	29	43	47	45	42	38	53	41	39	48	33	51
5	57	44	57	59	55	55	49	65	54	53	62	49	61
6	69	59	70	67	64	66	62	76	67	65	74	64	71
7	79	73	80	76	72	78	74	85	77	77	84	77	79
8	87	84	88	83	80	86	84	90	86	85	89	88	86
9	93	92	93	88	88	91	91	95	92	91	92	94	91
10	96	96	98	92	93	95	95	98	96	95	95	97	95
11	98	99	100	96	96	98	98	99	98	97	98	99	98
12	100	100	100	99	99	100	100	100	100	99	100	100	99

TABLE A-2. RAW SCORES AND PERCENTILES FOR THE U.S. NORM SAMPLE AND INTERNATIONAL SAMPLES—COLLABORATING

Raw Score	Percentile												
	English, North American	Chinese, Simplified	Chinese, Traditional	Danish	Dutch	French, Canadian	French, European	German	Portuguese, Brazilian	Portuguese, European	Spanish, European	Spanish, Latin American	Swedish
0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	0	0	1	2	0	1	1	0	1	2	0	3
2	3	1	2	3	8	2	4	4	3	2	6	4	9
3	7	5	8	8	20	9	8	12	10	6	15	13	21
4	15	15	18	17	36	20	20	24	24	14	30	30	36
5	26	31	34	31	53	36	38	39	44	27	47	50	52
6	41	52	57	47	71	55	57	58	65	48	65	69	68
7	58	73	77	63	87	73	74	74	82	69	82	86	81
8	74	89	91	79	96	88	90	87	93	87	93	93	90
9	87	97	97	92	99	96	98	96	98	97	97	97	96
10	95	99	99	97	99	99	99	99	100	99	99	99	99
11	99	100	100	100	100	100	100	100	100	100	100	100	100
12	100	100	100	100	100	100	100	100	100	100	100	100	100

TABLE A-3. RAW SCORES AND PERCENTILES FOR THE U.S. NORM SAMPLE AND INTERNATIONAL SAMPLES—COMPROMISING

Raw Score	Percentile												
	English, North American	Chinese, Simplified	Chinese, Traditional	Danish	Dutch	French, Canadian	French, European	German	Portuguese, Brazilian	Portuguese, European	Spanish, European	Spanish, Latin American	Swedish
0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	1	1	0	0	0	0	0	0	0	0
2	1	1	0	3	2	0	0	1	0	1	1	0	1
3	3	3	2	7	6	1	1	2	2	2	2	1	4
4	7	8	5	15	14	4	3	4	3	6	7	4	9
5	15	19	12	29	25	10	7	11	9	13	18	10	18
6	27	34	25	46	37	18	17	21	20	23	32	21	35
7	41	50	41	63	54	31	32	33	35	41	49	37	55
8	58	64	58	79	72	47	48	50	53	60	67	56	71
9	75	77	77	89	85	64	64	70	71	75	82	74	84
10	87	88	91	96	94	79	80	86	86	87	93	88	93
11	95	95	97	99	98	91	92	95	95	96	97	97	97
12	99	99	99	100	100	97	99	99	99	99	99	100	99

TABLE A-4. RAW SCORES AND PERCENTILES FOR THE U.S. NORM SAMPLE AND INTERNATIONAL SAMPLES—AVOIDING

Raw Score	Percentile												
	English, North American	Chinese, Simplified	Chinese, Traditional	Danish	Dutch	French, Canadian	French, European	German	Portuguese, Brazilian	Portuguese, European	Spanish, European	Spanish, Latin American	Swedish
0	1	0	0	0	0	0	0	0	0	0	0	0	0
1	2	0	0	1	0	1	0	1	0	0	0	0	0
2	6	1	2	2	1	5	3	4	1	3	2	2	1
3	12	6	6	4	3	11	9	8	4	7	5	7	3
4	22	19	17	10	8	22	21	15	13	15	10	13	8
5	34	39	34	20	16	39	39	27	28	27	20	23	18
6	49	63	54	35	28	58	60	44	45	46	34	37	33
7	65	81	73	52	43	73	76	61	64	62	51	55	48
8	78	92	87	69	59	85	86	78	80	75	66	75	63
9	88	98	95	83	75	93	94	89	90	86	82	88	77
10	95	100	99	92	90	97	97	94	95	94	93	95	89
11	98	100	100	97	97	99	99	98	98	98	97	98	96
12	100	100	100	99	99	100	100	100	100	99	100	100	99

TABLE A-5. RAW SCORES AND PERCENTILES FOR THE U.S. NORM SAMPLE AND INTERNATIONAL SAMPLES—ACCOMMODATING

Raw Score	Percentile												
	English, North American	Chinese, Simplified	Chinese, Traditional	Danish	Dutch	French, Canadian	French, European	German	Portuguese, Brazilian	Portuguese, European	Spanish, European	Spanish, Latin American	Swedish
0	0	0	0	0	0	0	1	0	0	1	0	0	0
1	2	1	2	1	1	2	4	2	0	4	1	1	1
2	7	3	4	3	4	6	10	5	4	13	4	6	3
3	16	9	9	11	10	15	18	10	12	25	9	16	8
4	30	17	18	22	20	29	29	19	26	37	17	31	15
5	46	31	32	38	32	45	43	32	42	52	29	46	26
6	62	51	47	53	45	61	58	49	59	68	43	62	40
7	76	70	61	67	61	76	74	65	76	79	59	75	56
8	87	87	76	80	75	86	87	78	87	88	73	86	72
9	94	95	88	89	87	93	95	88	95	95	85	94	85
10	98	98	95	96	96	98	99	95	98	98	94	98	95
11	100	99	99	99	99	99	100	99	100	100	98	100	99
12	100	100	100	100	100	100	100	100	100	100	100	100	100

TABLE A-6. TKI MODE PERCENTILE MEDIANS BY LANGUAGE USING THE U.S. NORM SAMPLE

TKI Mode	Percentile Median												
	English, North American	Chinese, Simplified	Chinese, Traditional	Danish	Dutch	French, Canadian	French, European	German	Portuguese, Brazilian	Portuguese, European	Spanish, European	Spanish, Latin American	Swedish
Competing	44	57	44	44	51	44	57	44	57	57	44	57	44
Collaborating	58	41	41	41	26	41	41	41	26	41	26	26	26
Compromising	58	41	58	27	41	58	58	58	58	41	41	58	41
Avoiding	49	34	49	65	65	49	49	49	49	49	65	65	65
Accommodating	46	62	62	62	62	46	46	62	62	46	62	46	76

Note: The median percentile indicates where half the scores are above and half the scores are below.

APPENDIX B: TUKEY'S HSD POST HOC ANALYSES

TABLE B-1. TUKEY'S HSD POST HOC ANALYSES FOR U.S. VS INTERNATIONAL LANGUAGES—COMPETING

Language	Competing			
	Mean Difference	Standard Error	<i>p</i>	Cohen's <i>d</i>
Chinese, Simplified	-0.729	.169	.001	-.28
Chinese, Traditional	0.050	.160	1.000	.02
Danish	0.011	.153	1.000	.00
Dutch	-0.171	.154	.997	-.06
French, Canadian	-0.184	.150	.991	-.07
French, European	-0.383	.159	.436	-.14
German	0.657	.158	.002	.24
Portuguese, Brazilian	-0.166	.160	.998	-.06
Portuguese, European	-0.315	.149	.656	-.11
Spanish, European	0.363	.156	.491	.13
Spanish, Latin American	-0.384	.153	.369	-.14
Swedish	0.324	.152	.642	.11

TABLE B-2. TUKEY'S HSD POST HOC ANALYSES FOR U.S. VS INTERNATIONAL LANGUAGES—COLLABORATING

Language	Collaborating			
	Mean Difference	Standard Error	<i>p</i>	Cohen's <i>d</i>
Chinese, Simplified	0.605	.127	.000	.31
Chinese, Traditional	0.822	.121	.000	.41
Danish	0.351	.115	.107	.16
Dutch	1.700	.116	.000	.83
French, Canadian	0.761	.113	.000	.37
French, European	0.878	.120	.000	.43
German	0.919	.118	.000	.43
Portuguese, Brazilian	1.180	.120	.000	.59
Portuguese, European	0.475	.112	.002	.24
Spanish, European	1.323	.117	.000	.63
Spanish, Latin American	1.386	.115	.000	.68
Swedish	1.550	.114	.000	.71

**TABLE B-3. TUKEY'S HSD POST HOC ANALYSES FOR
U.S. VS INTERNATIONAL LANGUAGES—COMPROMISING**

Language	Compromising			
	Mean Difference	Standard Error	<i>p</i>	Cohen's <i>d</i>
Chinese, Simplified	0.317	.130	.418	.14
Chinese, Traditional	-0.009	.123	1.000	.00
Danish	1.194	.118	.000	.55
Dutch	0.786	.118	.000	.36
French, Canadian	-0.654	.115	.000	-.30
French, European	-0.646	.122	.000	-.30
German	-0.372	.121	.102	-.17
Portuguese, Brazilian	-0.341	.123	.212	-.16
Portuguese, European	-0.041	.115	1.000	-.02
Spanish, European	0.393	.120	.055	.18
Spanish, Latin American	-0.216	.118	.831	-.10
Swedish	0.588	.117	.000	.28

**TABLE B-4. TUKEY'S HSD POST HOC ANALYSES FOR
U.S. VS INTERNATIONAL LANGUAGES—AVOIDING**

Language	Avoiding			
	Mean Difference	Standard Error	<i>p</i>	Cohen's <i>d</i>
Chinese, Simplified	0.540	.140	.008	.26
Chinese, Traditional	0.221	.133	.908	.10
Danish	-0.803	.127	.000	-.35
Dutch	-1.258	.128	.000	-.55
French, Canadian	0.367	.124	.137	.16
French, European	0.385	.132	.151	.17
German	-0.258	.131	.750	-.11
Portuguese, Brazilian	-0.266	.132	.725	-.12
Portuguese, European	-0.331	.124	.266	-.14
Spanish, European	-0.863	.129	.000	-.37
Spanish, Latin American	-0.507	.127	.004	-.22
Swedish	-1.090	.126	.000	-.47

**TABLE B-5. TUKEY'S HSD POST HOC ANALYSES FOR
U.S. VS INTERNATIONAL LANGUAGES—ACCOMMODATING**

Language	Accommodating			
	Mean Difference	Standard Error	<i>p</i>	Cohen's <i>d</i>
Chinese, Simplified	-0.539	.135	.005	-.26
Chinese, Traditional	-0.848	.129	.000	-.37
Danish	-0.555	.123	.000	-.25
Dutch	-0.849	.123	.000	-.37
French, Canadian	-0.055	.120	1.000	-.02
French, European	0.029	.127	1.000	.01
German	-0.723	.126	.000	-.32
Portuguese, Brazilian	-0.133	.128	.998	-.06
Portuguese, European	0.449	.120	.011	.19
Spanish, European	-1.007	.125	.000	-.44
Spanish, Latin American	0.003	.123	1.000	.00
Swedish	-1.154	.122	.000	-.51

APPENDIX C: TKI Mode Mean Score Differences Between Men and Women

TABLE C-1. MEAN DIFFERENCES BETWEEN MEN AND WOMEN IN THE INTERNATIONAL SAMPLES BY LANGUAGE—COMPETING

Language	Competing						df	t	p	Cohen's d
	Men			Women						
	N	M	SD	N	M	SD				
Chinese, Simplified	137	5.42	2.37	148	5.41	2.56	283.00	-0.06	.95	-.01
Chinese, Traditional	153	4.90	2.58	165	4.90	2.58	316.00	-1.72	.09	-.19
Danish	170	5.06	3.34	180	4.31	2.98	338.27 ^a	-2.24	.03	-.24
Dutch	171	5.02	3.48	175	4.70	2.95	332.77 ^a	-0.92	.36	-.10
French, Canadian	181	4.99	2.69	186	4.75	2.80	365.00	-0.86	.39	-.09
French, European	157	5.36	2.73	167	4.80	2.81	322.00	-1.82	.07	-.20
German	159	4.27	2.74	171	3.80	2.73	328.00	-1.56	.12	-.17
Portuguese, Brazilian	152	4.91	2.87	169	4.79	2.62	319.00	-0.40	.69	-.04
Portuguese, European	187	5.19	2.95	182	4.80	2.58	362.83 ^a	-1.35	.18	-.14
Spanish, European	159	4.67	2.94	180	4.02	2.78	337.00	-2.09	.04	-.23
Spanish, Latin American	179	5.27	2.55	172	4.85	2.42	349.00	-1.58	.12	-.17
Swedish	171	4.75	2.98	184	3.99	3.03	353.00	-2.38	.02	-.25

Note: df = degrees of freedom; t = t-value; p = p-value; ^a = unequal variance.

TABLE C-2. MEAN DIFFERENCES BETWEEN MEN AND WOMEN BY LANGUAGE—COLLABORATING

Language	Collaborating						df	t	p	Cohen's d
	Men			Women						
	N	M	SD	N	M	SD				
Chinese, Simplified	137	5.85	1.68	148	5.91	1.71	283.00	0.26	.80	.03
Chinese, Traditional	153	5.66	1.77	165	5.66	1.77	316.00	0.03	.97	.00
Danish	170	6.11	2.08	180	6.16	2.13	348.00	0.25	.81	.03
Dutch	171	4.87	1.95	175	4.71	1.86	344.00	-0.77	.44	-.08
French, Canadian	181	5.75	1.89	186	5.70	1.89	365.00	-0.27	.79	-.03
French, European	157	5.54	1.91	167	5.68	1.83	322.00	0.68	.50	.08
German	159	5.50	2.01	171	5.63	2.03	328.00	0.60	.55	.07
Portuguese, Brazilian	152	5.23	1.74	169	5.37	1.75	319.00	0.73	.47	.08
Portuguese, European	187	6.22	1.73	182	5.80	1.80	367.00	-2.30	.02	-.24
Spanish, European	159	5.26	2.01	180	5.07	1.95	337.00	-0.89	.37	-.10
Spanish, Latin American	179	5.31	1.88	172	4.88	1.75	349.00	-2.24	.03	-.24
Swedish	171	4.82	2.18	184	5.04	2.19	353.00	0.97	.33	.10

Note: df = degrees of freedom; t = t-value; p = p-value; ^a = unequal variance.

TABLE C-3. MEAN DIFFERENCES BETWEEN MEN AND WOMEN BY LANGUAGE—COMPROMISING

Language	Compromising						df	t	p	Cohen's d
	Men			Women						
	N	M	SD	N	M	SD				
Chinese, Simplified	137	7.00	2.17	148	7.20	2.31	283.00	0.76	.44	.09
Chinese, Traditional	153	7.27	2.07	165	7.27	2.07	316.00	1.42	.16	.16
Danish	170	6.12	2.33	180	6.33	1.95	330.17 ^a	0.94	.35	.10
Dutch	171	6.54	2.30	175	6.73	2.17	344.00	0.76	.45	.08
French, Canadian	181	7.93	2.13	186	8.22	2.19	365.00	1.25	.21	.13
French, European	157	7.92	2.04	167	8.20	2.06	322.00	1.23	.22	.14
German	159	7.55	2.06	171	8.02	2.04	328.00	2.11	.04	.23
Portuguese, Brazilian	152	7.64	1.94	169	7.87	2.00	319.00	1.02	.31	.11
Portuguese, European	187	7.26	2.02	182	7.67	2.14	367.00	1.88	.06	.20
Spanish, European	159	7.01	1.93	180	7.04	2.18	337.00	0.14	.89	.02
Spanish, Latin American	179	7.64	1.95	172	7.63	1.91	349.00	-0.04	.97	.00
Swedish	171	6.73	2.13	184	6.93	2.03	353.00	0.90	.37	.10

Note: df = degrees of freedom; t = t-value; p = p-value; ^a = unequal variance.

TABLE C-4. MEAN DIFFERENCES BETWEEN MEN AND WOMEN BY LANGUAGE—AVOIDING

Language	Avoiding						df	t	p	Cohen's d
	Men			Women						
	N	M	SD	N	M	SD				
Chinese, Simplified	137	5.38	1.60	148	5.63	1.61	283.00	1.30	.19	.16
Chinese, Traditional	153	5.87	1.78	165	5.87	1.78	316.00	-0.40	.69	-.04
Danish	170	6.84	2.12	180	6.87	2.27	348.00	0.13	.89	.01
Dutch	171	7.36	2.17	175	7.26	2.17	344.00	-0.43	.67	-.05
French, Canadian	181	5.65	2.15	186	5.72	2.16	365.00	0.31	.76	.03
French, European	157	5.50	1.96	167	5.81	2.06	322.00	1.39	.17	.15
German	159	6.26	2.28	171	6.35	2.19	328.00	0.38	.71	.04
Portuguese, Brazilian	152	6.51	2.08	169	6.14	1.97	319.00	-1.61	.11	-.18
Portuguese, European	187	6.31	2.33	182	6.45	2.19	367.00	0.60	.55	.06
Spanish, European	159	6.78	2.18	180	7.03	2.18	337.00	1.05	.30	.11
Spanish, Latin American	179	6.29	2.14	172	6.83	2.07	349.00	2.40	.02	.26
Swedish	171	7.14	2.19	184	7.14	2.24	353.00	-0.02	.98	.00

Note: df = degrees of freedom; t = t-value; p = p-value; ^a = unequal variance.

TABLE C-5. MEAN DIFFERENCES BETWEEN MEN AND WOMEN BY LANGUAGE—ACCOMMODATING

Language	Accommodating						df	t	p	Cohen's d
	Men			Women						
	N	M	SD	N	M	SD				
Chinese, Simplified	137	6.12	1.84	148	5.70	2.02	283.00	-1.83	.07	-.22
Chinese, Traditional	153	6.10	2.35	165	6.10	2.35	316.00	0.81	.42	.09
Danish	170	5.71	2.20	180	6.11	2.36	348.00	1.61	.11	.17
Dutch	171	6.02	2.44	175	6.39	2.29	344.00	1.44	.15	.15
French, Canadian	181	5.47	2.09	186	5.36	2.41	360.06 ^a	-0.46	.64	-.05
French, European	157	5.38	2.41	167	5.29	2.23	322.00	-0.34	.73	-.04
German	159	6.22	2.35	171	5.95	2.32	328.00	-1.04	.30	-.11
Portuguese, Brazilian	152	5.46	2.04	169	5.52	2.06	319.00	0.26	.79	.03
Portuguese, European	187	4.76	2.33	182	5.07	2.47	367.00	1.23	.22	.13
Spanish, European	159	6.11	2.46	180	6.59	2.27	337.00	1.90	.06	.21
Spanish, Latin American	179	5.21	2.19	172	5.51	2.24	349.00	1.29	.20	.14
Swedish	171	6.32	2.22	184	6.69	2.31	353.00	1.53	.13	.16

Note: df = degrees of freedom; t = t-value; p = p-value; ^a = unequal variance.