

EVALUATING MEASUREMENT EQUIVALENCE AND TRANSLATION EFFECTIVENESS OF A CUSTOMER ENGAGEMENT INSTRUMENT ACROSS NATIONAL CULTURES AND TYPES OF CUSTOMERS

Dan YU

Yongwei YANG

Gallup Organization

May 21, 2014

2014 Modern Modeling Methods Conference

WHY NEED TO STUDY CROSS GROUPS MEASUREMENT EQUIVALENCE OF CUSTOMER INSTRUMENT?

Understanding customers' perceptions and managing customer relationships are critical to enterprise success.

Many organizations today have markets and conduct business across countries.

NEEDS

Use the same survey to gather customer data from these multiple cultures, then measure, compare and manage *customer engagement* in international settings.

National culture affects consumers' attitudes and how they respond to survey questions

Translation of survey questions may introduce construct irrelevant variance

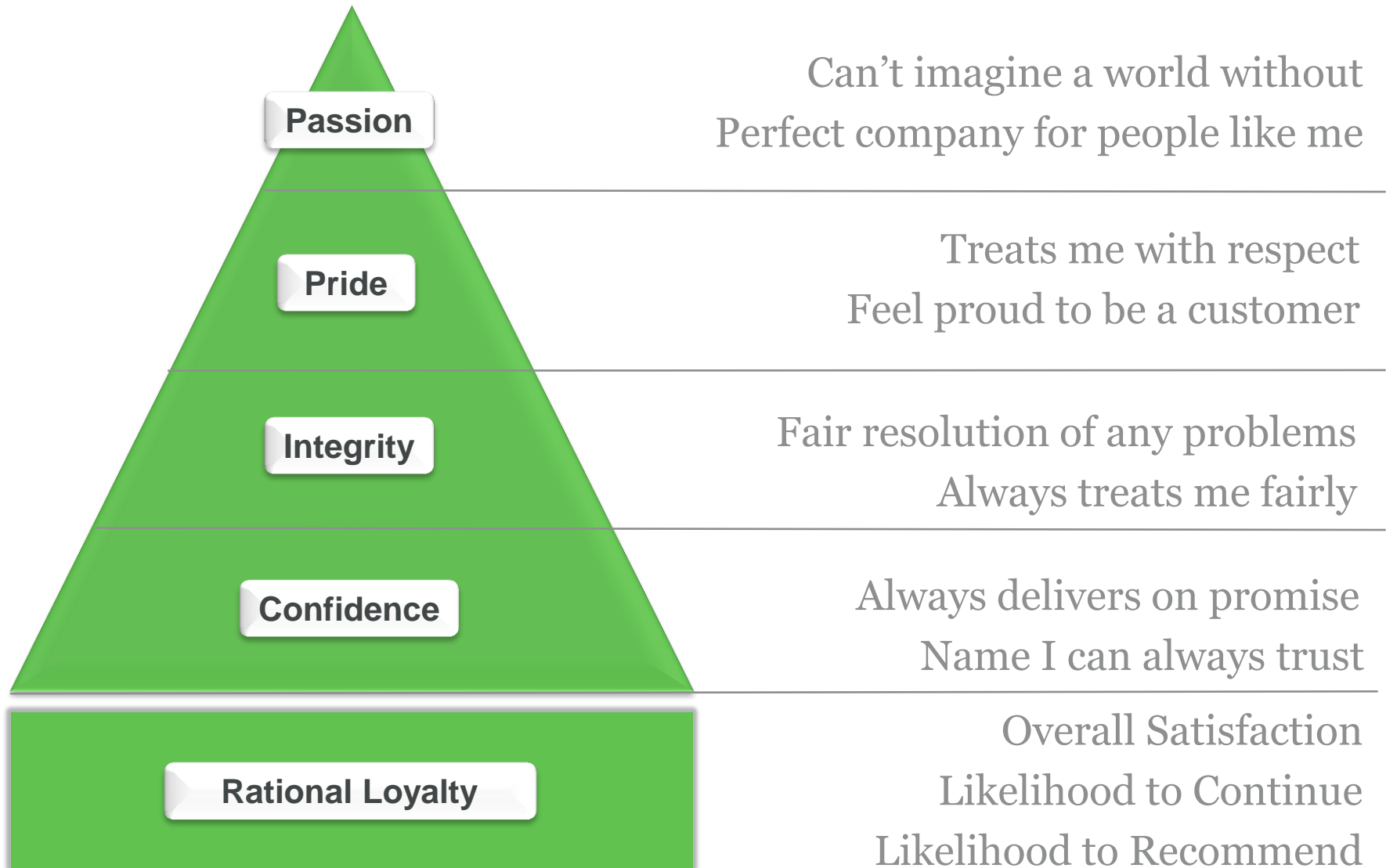
Customer survey differences may exist between B2B and B2C customers

All in all, for valid comparisons to be made among individual or business customers across cultures, the measurement properties of the evaluation survey must be consistent across these groups.

HOW TO STUDY CROSS GROUPS MEASUREMENT EQUIVALENCE OF CUSTOMER INSTRUMENT??

- A range of statically methods exist for evaluating and establishing construct equivalence such as multiple-group confirmatory factor analysis(CFA), weighted multidimensional scaling(WMDS), MIMIC modeling, latent class models and mixture models etc.
- **In this study, we use WMDS for construct equivalence.**
 - Unlike CFA, MDS does not require pre-specification of the underlying structure.
 - MDS can readily model ordinal data
 - MDS is available in basic statically software
 - MDS allows for easy visual representations of the construct differences which make the findings intuitively accessible to non-statistical audience.
- **In this study, we use DIF to evaluate item level equivalence.**
 - Three regression-based approaches: ordinal logistic, linear and ANCOVA and then compare and combine the results of all three methods.
 - Three types of analyses to answer three questions.
 - ✓ Comparing customer types: Can we measure and compare scores btw. B2B and B2C customers?
 - ✓ Cultural difference: Can we measure and compare scores btw. focal Language-Country and English-U.S.?
 - ✓ Cultural uniqueness: Can we measure and compare scores btw. focal Language-Country and the rest?

METHOD-CUSTOMER INSTRUMENT

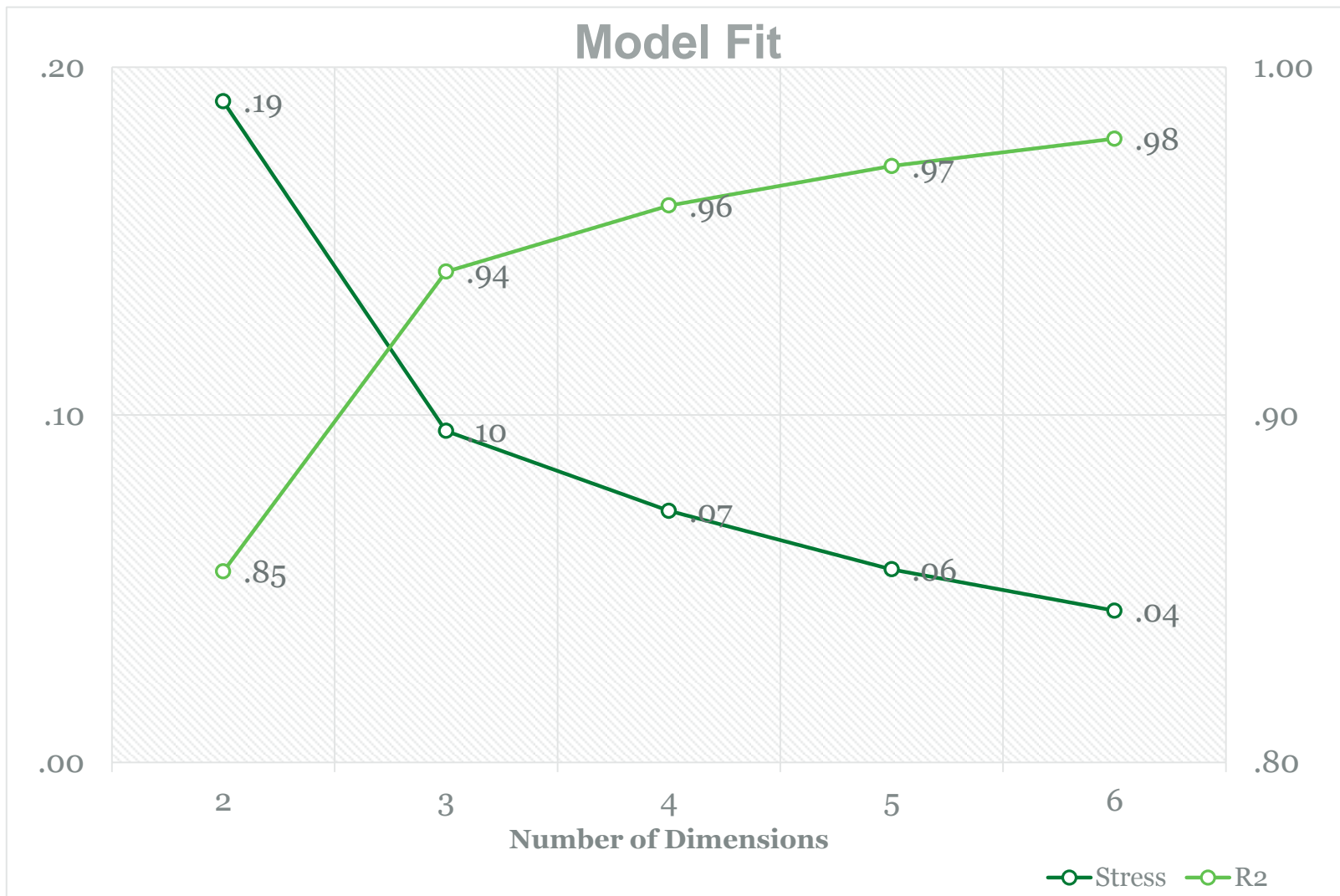


METHOD—CULTURE GROUPS AND CUSTOMER BUSINESS TYPES; ANALYSIS SAMPLES

Table 1. Sample sizes in MDS and DIF analysis.

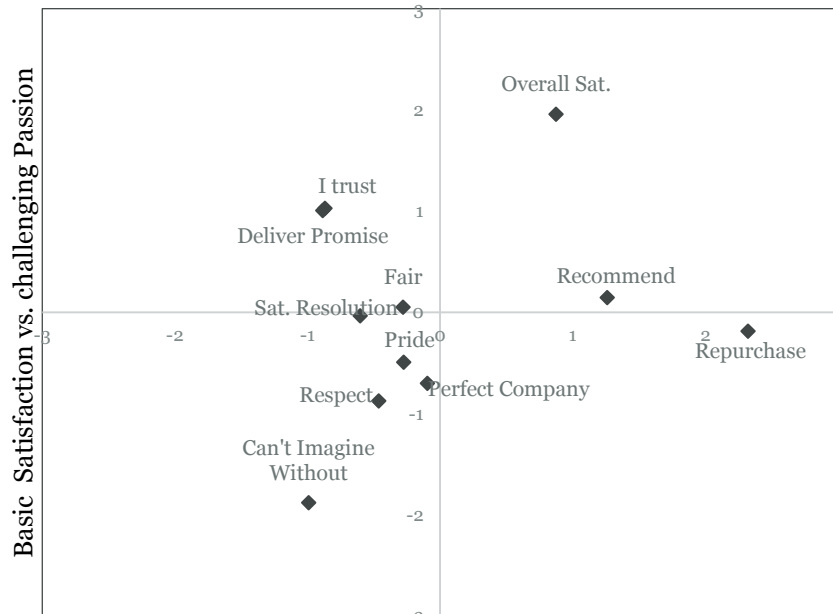
Culture Group	Customer Type	Sample 2 for MDS and DIF Analysis	Sample Size for Customer Type DIF Analysis	Sample Size for “Versus-U.S.” Cross-cultural DIF Analysis (country: U.S.)	Sample Size for “Versus-Other” Cross-cultural DIF Analysis: Versus-Other Analysis (country : other total / per country)
U.S.-English	B2C	1,000	1,000	--	1,000 : 1,008 / 72
U.S.-English	B2B	1,000	1,000	--	1,000 : 1,008 / 72
Panama-Spanish	B2C	815	572	815 : 815	815 : 826 / 59
Panama-Spanish	B2B	572	572	572: 572	572 : 574 / 41
Venezuela-Spanish	B2C	1,000	1,000	1,000 : 1,000	1,000 : 1,008 / 72
Venezuela-Spanish	B2B	1,000	1,000	1,000 : 1,000	1,000 : 1,008 / 72
UK-English	B2C	1,000	1,000	1,000 : 1,000	1,000 : 1,008 / 72
UK-English	B2B	1,000	1,000	1,000 : 1,000	1,000 : 1,008 / 72
France-French	B2C	1,000	219	1,000 : 1,000	1,000 : 1,008 / 72
France-French	B2B	219	219	219 : 219	219 : 224 / 16
Netherland-Dutch	B2C	1,000	264	1,000 : 1,000	1,000 : 1,008 / 72
Netherland-Dutch	B2B	264	264	264 : 264	264 : 266 / 19
Germany-German	B2C	1,000	1,000	1,000 : 1,000	1,000 : 1,008 / 72
Germany-German	B2B	1,000	1,000	1,000 : 1,000	1,000 : 1,008 / 72
Italy-Italian	B2C	1,000	132	1,000 : 1,000	1,000 : 1,008 / 72
Italy-Italian	B2B	132	132	132 : 132	132 : 140 / 10
Hungary-Hungarian	B2C	1,000	1,000	1,000 : 1,000	1,000 : 1,008 / 72
Hungary-Hungarian	B2B	1,000	1,000	1,000 : 1,000	1,000 : 1,008 / 72
India-English	B2C	1,000	1,000	1,000 : 1,000	1,000 : 1,008 / 72
India-English	B2B	1,000	1,000	1,000 : 1,000	1,000 : 1,008 / 72
India-Native Languages	B2C	1,000	1,000	1,000 : 1,000	1,000 : 1,008 / 72
India-Native Languages	B2B	1,000	1,000	1,000 : 1,000	1,000 : 1,008 / 72
Mainland China-Mandarin	B2C	1,000	1,000	1,000 : 1,000	1,000 : 1,008 / 72
Mainland China- Mandarin	B2B	1,000	1,000	1,000 : 1,000	1,000 : 1,008 / 72
Hong Kong -Cantonese	B2C	1,000	224	1,000 : 1,000	1,000 : 1,008 / 72
Hong Kong-Cantonese	B2B	224	224	224 : 224	224 : 224 / 16
Japan-Japanese	B2C	1,000	1,000	1,000 : 1,000	1,000 : 1,008 / 72
Japan-Japanese	B2B	1,000	1,000	1,000 : 1,000	1,000 : 1,008 / 72
Thailand-Thai	B2C	1,000	1,000	1,000 : 1,000	1,000 : 1,008 / 72
Thailand-Thai	B2B	1,000	1,000	1,000 : 1,000	1,000 : 1,008 / 72

WMDS ANALYSIS-DETERMINE NO. OF DIMENSIONS



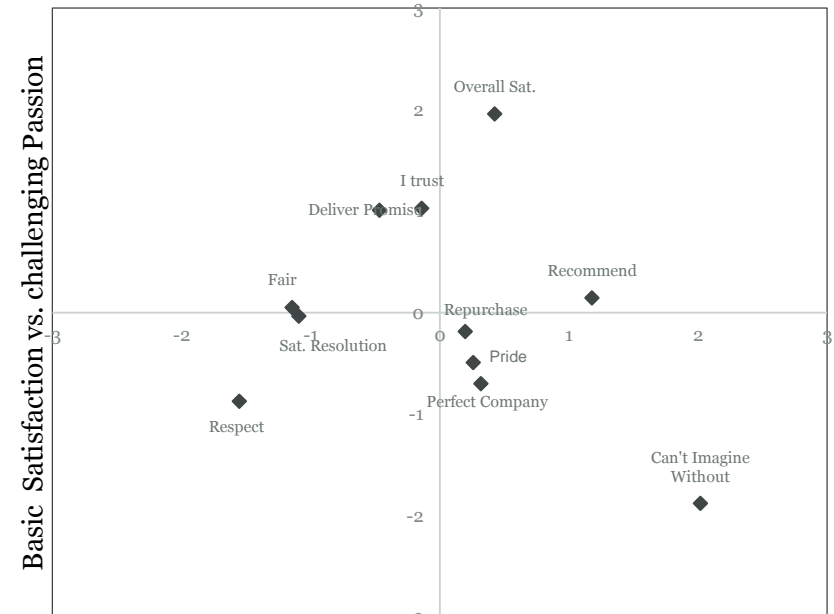
WMDS ANALYSIS-INTERPRETING THE 3 DIMENSIONS

Stimulus Space: Dimension 1 by Dimension 2
U.S.-ENG-B2C Sample (n=10,000)



Rational loyalty vs. Emotional attachment

Stimulus Space: Dimension 2 by Dimension 3
U.S.-ENG-B2C Sample (n=10,000)



Broad brand perception vs. tangible specific interaction

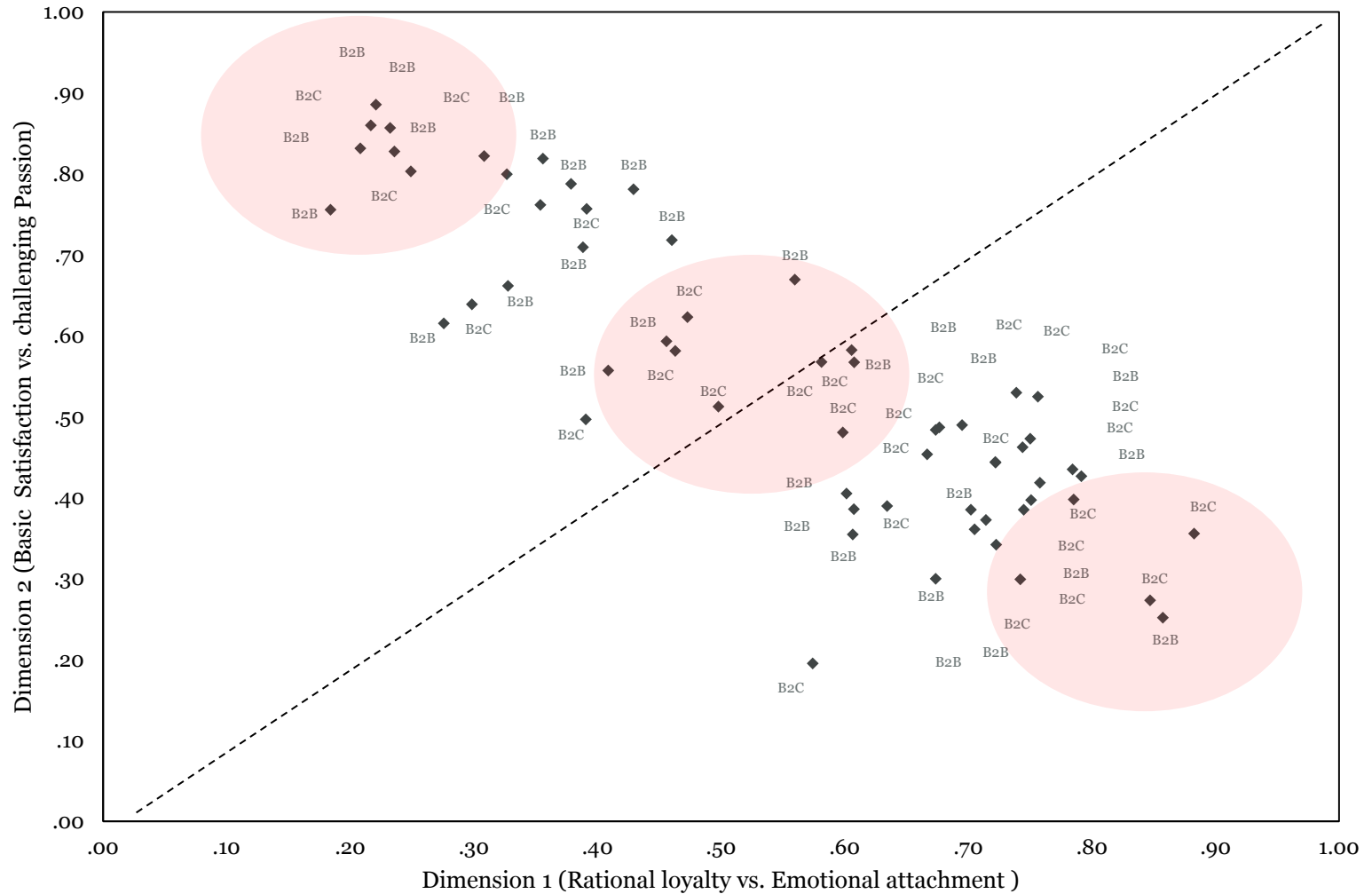
Dimension 1: Rational loyalty vs. Emotional attachment

Dimension 2: Basic satisfaction vs. the most challenging passion

Dimension 3: Broad brand perception vs. tangible specific interaction

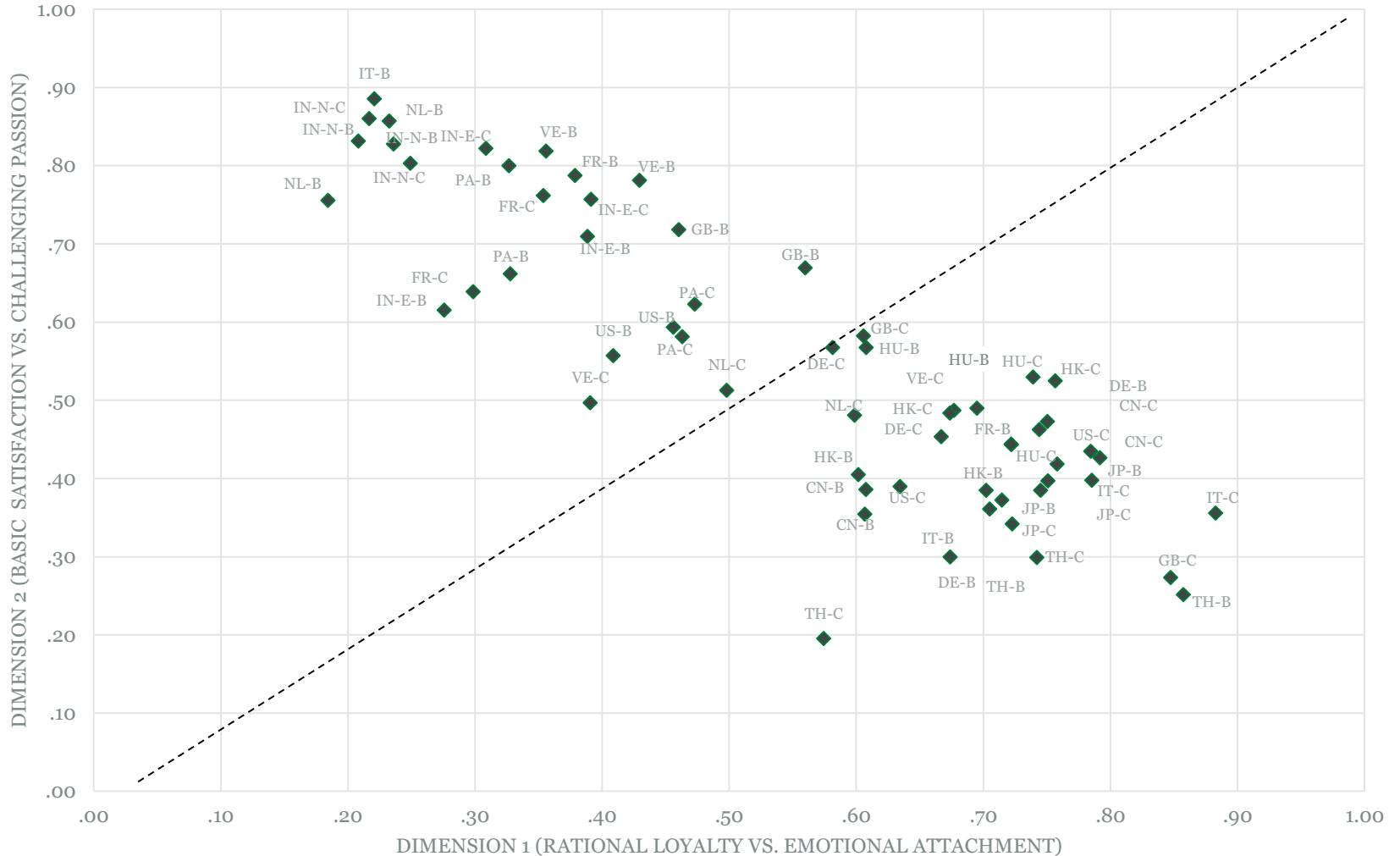
WMDS ANALYSIS—STRUCTURE CONSISTENCY BETWEEN B2B AND B2C GROUPS

Subject Space (Labeled by Customer Type):
Dimension 1 by Dimension 2



WMDS ANALYSIS--STRUCTURE CONSISTENCY AMONG COUNTRY-LANGUAGE GROUPS

SUBJECT SPACE (LABELED BY COUNTRY AND BUSINESS TYPE):
DIMENSION 1 BY DIMENSION 2



DIF(ORDINAL LOGISTIC REGRESSION) ANALYSIS RESULTS FOR CUSTOMER TYPE

Culture Group	CE01. Overall Product/Service Satisfaction			CE03. Likelihood to Recommend a Friend or Associate			CE06. Always treats me fairly			CE07. If a problem arises, I can always count on fair and satisfactory resolution			CE08. I feel proud to be a customer			CE09. Always treats me with respect			CE11. I can't imagine a world without			
	Overall	Uniform	Non-Uniform	Overall	Uniform	Non-Uniform	Overall	Uniform	Non-Uniform	Overall	Uniform	Non-Uniform	Overall	Uniform	Non-Uniform	Overall	Uniform	Non-Uniform	Overall	Uniform	Non-Uniform	
U.S.-English																						
Panama-Spanish																						
Venezuela-Spanish																						
UK-English																M	M			M	M	
France-French	M	M		M	M																	
Netherland-Dutch																						
Germany-German																						
Italy-Italian							M	M		M	M					L	L			L	M	
Hungary-Hungarian																						
India-English																						
India-Native Languages																						
Mainland China-Mandarin													M	M								
Hong Kong-Cantonese				L	L																	
Japan-Japanese																						
Thailand-Thai																						

Note: DIF cutoff points: $S < 0.035$; $0.035 \leq M \leq 0.069$; $0.070 \leq L$.

DIF(ORDINAL LOGISTIC REGRESSION) ANALYSIS RESULTS

B2B vs B2C in each culture (165 regressions)

- No large or moderate non-uniform DIF
- 3 large overall DIF
- 7 moderate overall DIF

Focal culture vs. English-USA (308 regressions)

- No large or moderate non-uniform DIF
- 0 large overall DIF
- 6 moderate overall DIF

Focal culture vs. combination of the rest (330 regressions)

- No large or moderate non-uniform DIF
- 3 large overall DIF
- 10 moderate overall DIF

Summary of DIF Findings

1. Both the no. of questions flagged for DIF and their effect sizes are quite small.
2. The DIF effects are primarily due to uniform DIF, rather than non-uniform DIF (the serious indicator of systematic difference).
3. The small number of flagged DIFs are scattered, rather than always appearing on the same question or the same group studied.

Conclusion

In general, the 11 questions in the instrument function similarly across culture groups and customer types.

CONCLUDING REMARKS—APPLICATION OF INSTRUMENTS IN PRACTICE

- For international businesses, it is very important to establish measurement equivalence before cross-cultural application of any customer metric.
- Overall, this study shows practitioners can apply this customer metric in the 14 national cultures for both B2B and B2C customers with a high degree of confidence that it measures the same thing in the same way.
- Without the evidence about measurement equivalence, it is advisable to adopt a conservative approach of only comparing the scores of one company or business unit with peers within the same national culture.
- Careful translation-adaption processes can be effective for preserving construct equivalence across multiple-language versions of a survey. Even so, translation is not a perfect science and the DIF analysis can help identify which question or language needs improvement.

CONCLUDING REMARKS--METHOD

- This study illustrates the utility of WMDS for assessing construct equivalence when underlying structure is not clearly known, the response data may not be ratio or interval, and when communicating the findings to non-statistical audience is important.
- The study also shows the significant value of using DIF methods for assessing cultural, language, or translation effects.
- By presenting these methods and also discussing the conditions and rationales for using these techniques, we encourage more researchers and business practitioners to use these statistical methods to evaluate ME/I in their works and more survey methodologists to study the utility and limitations of these methods.

Q&A

COPYRIGHT STANDARDS

This document contains proprietary research, copyrighted materials, and literary property of Gallup, Inc. It is for the guidance of your company only and is not to be copied, quoted, published, or divulged to others outside of your organization. Gallup® and Gallup, Inc. are trademarks of Gallup, Inc. All other trademarks are the property of their respective owners.

This document is of great value to both your organization and Gallup, Inc. Accordingly, international and domestic laws and penalties guaranteeing patent, copyright, trademark, and trade secret protection protect the ideas, concepts, and recommendations related within this document.

No changes may be made to this document without the express written permission of Gallup, Inc.