Developing a Model to Conceptualize the Results of Comparative Pretesting:
‘A Presentation’

Gordon Willis, Ph.D.
Applied Research Program
Division of Cancer Control and Population Sciences
National Cancer Institute
Willisg@mail.nih.gov
Objective: Characterize the findings from pretesting of multicultural surveys

“Error Models” of the survey response process already exist:

(1) Tourangeau (1983): 4-Stage Cognitive Model
• Comprehension
• Retrieval
• Decision/Judgment
• Response


(3) Q-BANK (Miller, Maitland, et al.): Response Error Indicator Codes (similar to QAS)
Why do we need a new model for cross-cultural/multilingual surveys?

• Because cross-cultural investigations include new, or exacerbated, elements, especially when translated
  – *Translation* step adds complexity
  – *Survey response is NOT just a cognitive enterprise, but a socio-cultural one*
  – *As such, there is non-ignorable variance in the ways that groups respond to survey questions*

• Empirical cross-cultural investigations have led to a somewhat novel approach to conceptualizing error →
1) Translation problems
2) Problems of Cultural Adaptation
3) Generic problems of question design
TCG Model of Cross-Cultural Error: Translation Problems

1) $T = \textit{Error or difficulty}$ in conversion of meaning of word, phrase, question ("Surface structure")

- **Error**: The term “excellent” is translated into “God-like” in Chinese.

- **Difficulty**: In general, vague quantifiers like “excellent” – “poor” present a challenge to mapping into other languages.
2) \( C = \text{Underlying conceptualization} \) (e.g., “Deep structure”) is not equivalent in different languages:

- Mexican (4-meal) pattern: 3-meal “Breakfast/lunch/dinner” pattern doesn’t apply.

- Questions about tar/nicotine level on cigarette brand smoked the most don’t account for the fact that Korean and Chinese brands don’t have these specified.
3) G = Problems that affect everyone

- These might map back to the Tourangeau, QAS, or Q-BANK models
- There are common features that influence all groups:
  - “Vigorous” versus “Moderate” physical activity
  - Vague terms
  - How long is it before you smoke the first cigarette of the day -> responses in terms other than minutes/hours
A recent study (Berrigan et al., In prep) illustrates a common pattern: Generic problems are ubiquitous:

- **Physical Activity questions:**
  - T: 2%  
  - C: 8%  
  - G: 52%

- **Acculturation questions:**
  - T: 2%  
  - C: 26%  
  - G: 49%

So, many problems appear to be grounded in a “failure to communicate,” generally

(1) Translation errors: Translated questions are not functionally equivalent to source questions.

(2) Cultural issues: When the concepts being measured do not exist in a similar form, across countries.

(3) Source question errors: All or part of the source question has been poorly designed leading to problems which are found across all countries.

So far, a 1:1 match with TCG system ->
(4) Source question and its interaction with translation: When the question works well in the source questionnaire, but has features in its design which make translation difficult, leading to measurement problems.

Example: Translating “all, almost, most, some, a few, no” from English produces non-equivalence.

BUT: What is the difference between “Translation error” and “Source question and its interaction with translation”? 

Translation error = Translation from Source to Target is possible, but was done erroneously: **Translator Error**

Source question and its interaction with translation = Translation from Source to Target is very problematic in the first place: **Translation Difficulty / Problem**

- This seems reasonable – There are lots of terms that don’t translate easily

So, Fitzgerald et al. divide TGC “T” problems (Error or Difficulty) into two sub-components: (1) Error and (2) Difficulty

Translation error = Translation from Source to Target is possible, but was done erroneously: **Translator Error**

Source question and its interaction with translation = Translation from Source is very problematic in the first place: **Translation Difficulty / Problem**

- This seems reasonable — There are lots of terms that don’t translate easily: (German “Volk”)

So, Fitzgerald et al. divide TGC “T” problems (Error or Difficulty) into two sub-components: (1) Error and (2) Difficulty

Maybe these would be better labels?
Alternative View of Alternative Model of Cross-Cultural Error:

1) Translation error
   1a) Translator Error
   1b) Translation Difficulty

2) Cultural problems

3) Generic Problems
Overall: Good example of convergence of conceptualizations, with respect to a (presumably) non-trivial issue

Q: But, how are these models useful?

A: We take very different approaches to fixing problems, based on their source:

- If **Generic**: we fix up the source question
- If **Translator Error**, we have the translator fix this
- If **Translation Difficulty** OR a **Cultural** problem, we may have need to work harder, to obtain “input harmonization” such that everyone understands the concepts similarly