

# **Translating and adapting tests**



# Outline

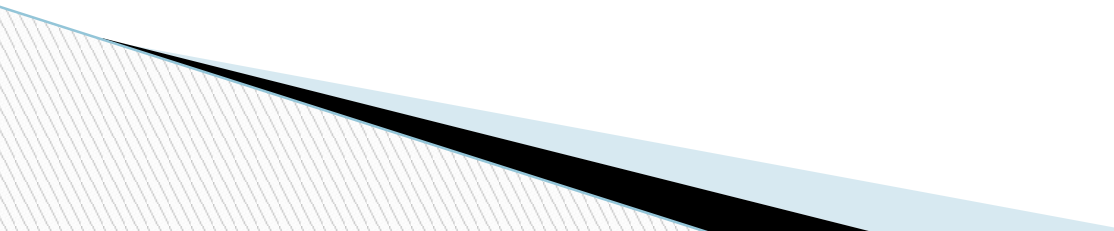
- Outline: What is a good translation?
- History: Two generations
- First generation
  - Focus on close translations
  - Translation - back translation
- Second generation: From translations to adaptations
  - Integration of perspectives
    - Linguistic
    - Cultural
    - Psychometric
  - Multiple tests of quality
- Conclusion

# What Does “Equivalent” Mean?

- Eusebius Hieronymus (St. Jerome, famous bible translator from Greek and Hebrew to Latin; ±347—419/420):
  - 2 types of translations: “words” and “meanings” (he favored the latter)

# Translatability

A psychological test/item is

- **Well translatable** if linguistic and psychological features yield the same translation
  - **Poorly translatable** if linguistic and psychological features do not entirely converge (e.g., translation of slang: meaning is translatable, but conciseness is lost)
  - **Non-translatable** if there is a complete or nearly complete nonoverlap (e.g., Jabberwocky)
- 

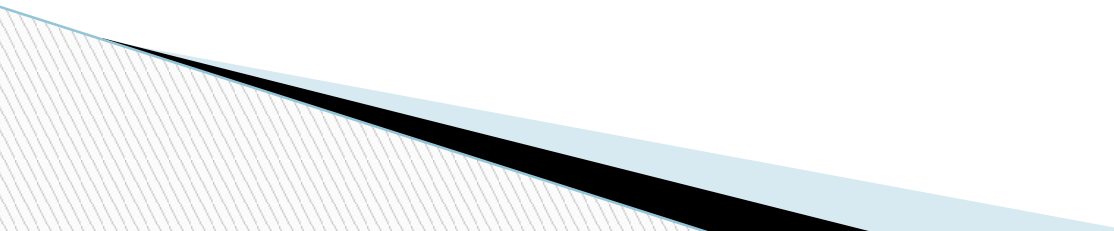
# Options

- ▣ **Adoption** (Close “literal” translation)
  - Advantage: maintains metric equivalence
  - Disadvantage: adequacy (too) readily assumed, should be demonstrated
- ▣ **Adaptation (changing contents of one or more items so as to increase cultural appropriateness)**
  - Advantage: more flexible, more tailored to the context
  - Disadvantage: fewer statistical techniques available to compare scores across cultures
- ▣ **Assembly (composing a new instrument)**
  - Advantage: very flexible
  - Disadvantage: almost no comparability maintained

# Comparative Studies: The First Generation

- Preparing Stimulus Materials
  - Emphasis on
    - Use of the same instrument in various places
    - Close / literal translations
- Quality assurance (Brislin, 1970)
  - Forward translation and independent back translation as quality check

# Problems with the Approach

- Emphasis on translations of existing materials
    - Capitalizes on equivalence, not on how the new cultural context could be different
    - Instruments developed in a Western context may not provide a good measure in a new context
- 

# Translation – Back Translation

Journal of Cross-Cultural Psychology  
Vol. 1, No. 3, September 1970, pp. 185-216.

BACK-TRANSLATION FOR CROSS-CULTURAL RESEARCH<sup>1</sup>

RICHARD W. BRISLIN<sup>2</sup>  
The Pennsylvania State University

- Popular way of checking quality of translations
- Assumption
  - Quality of translation is adequate if back translation is more or less similar to original
- Advantage
  - No knowledge of target language is needed



# Translation – Back Translation

## ▣ Problems

- Can produce unnatural language
- No check on readability, natural flow
- Assumes high level of translatability
- Easy to cheat when translators know that their text will be back translated

# Initiatives to Enhance Quality

- ▣ In 1992 the International Test Commission (ITC) began a project to prepare guidelines for translating and adapting tests and psychological instruments, and establishing score equivalence across language and/or cultural groups
- ▣ Associations involved
  - European Association of Psychological Assessment
  - European Test Publishers Group
  - International Association for Cross-Cultural Psychology
  - International Association of Applied Psychology
  - International Association for the Evaluation of Educational Achievement
  - International Language Testing Association
  - International Union of Psychological Science.

# Core Elements of New Approach

- Translation □ adaptation
- Multiple types of expertise required
- Adaptation affects the whole research process
  - Requires both qualitative and quantitative aspects
  - Relevance of quality assurance as chain management

# Quality Assessment Procedures (after Harkness, 2003)

Translation stage

Pretesting stage  
(qualitative)

Pretesting or actual  
administration  
(quantitative)

forwa  
kt  
ende  
Cher  
Indep  
Y.e.g.,  
ass  
read  
p  
and  
Mixe  
on  
(s)  
ation  
d  
Gom  
(forw  
sach  
qual  
bilin  
and  
al  
an  
mon  
trist  
ion  
bi  
d  
an  
e  
nc  
s  
le  
focu  
ds  
va  
un

# Taxonomy of Adaptations

Need for adaptation can be due to:

## Construct equivalence

- Similarity of construct in source and target culture

## Cultural equivalence

- Norms about interaction (modes of address)
- “Cultural fact sheet”

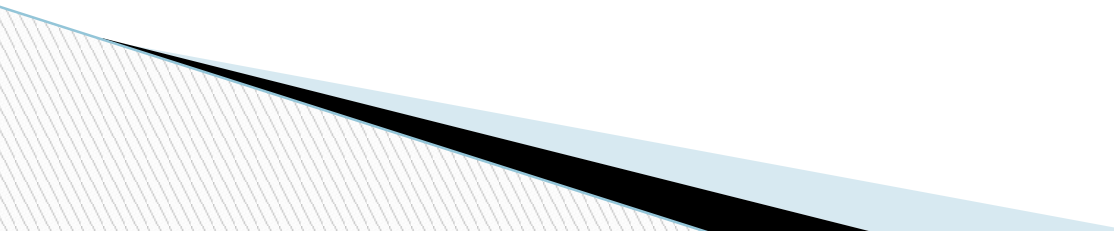
## Linguistic equivalence

- Translation accuracy: Retention of denotation and connotation

## Measurement equivalence

- Retention of psychometric features (response styles)
- Similarity of factors measured by a test and comparability of scores

# 1. Example Construct-Driven

- Problems with the behaviors or attitudes associated with the construct or with communication norms pertaining to these behaviors or attitudes
  - Usage of somatic and psychological symptoms in depression inventories
    - Differential norms in allowance to express psychological symptoms across cultures
- 

## 2. Example Culture-Driven

Example: 'Burglar' (Picture Arrangement; adapted for use in low-SES children in Bangalore, India by Malda et al., 2008)



### Problems:

1. Unclear whether the burglar was getting in or getting out;
2. Man not recognized as burglar;
3. Window was not recognized (vertically moving windows are uncommon in India)

# 3. Example of Language-Driven Adaptation

- Example: Do you often feel distressed?
- Translation to Dutch:
  - “Distressed” does not have an equivalent word in Dutch
  - Possible solutions
    - Composite of different emotions in Dutch; ask for frequency of composite (“how often do you feel X and Y?”). Problem: composite may not be recognizable
    - Choose a single emotion that is as close as possible; problem: change of item content if no close match can be found
    - Describe the emotion in the item (e.g., vignette); problem: may require a similar description in English original
  - Need to check adequacy of chosen solution in statistical analysis
  - Combination of judgmental and statistical evidence crucial in instruments that are more difficult to translate/adapt



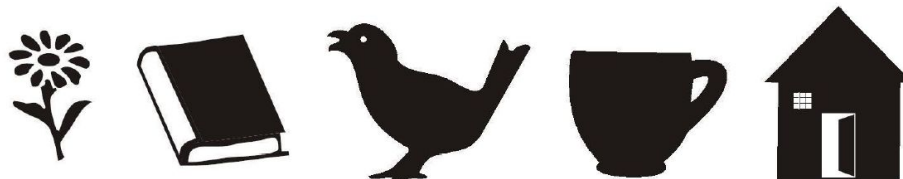
# 4. Example of Measurement-Driven Adaptation (Unfamiliarity)

- ❑ Kaufman ABC used in Bangalore (Kannada-speaking children)
- ❑ Adaptation of words in subtest *Word Order* based on:
  - Unfamiliarity and ambiguity of objects and words
  - Number of syllables

Original version ❑



Kannada version ❑



# 4. Example of Measurement-Driven Adaptation (Unfamiliarity)

Original version



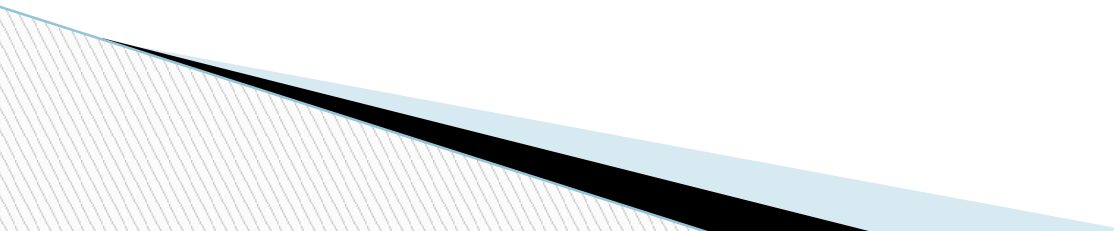
Kannada version



*Problems:*

- (1) Key was often called ice cream;
- (2) English word "key" was often used, which is too short (monosyllabic)

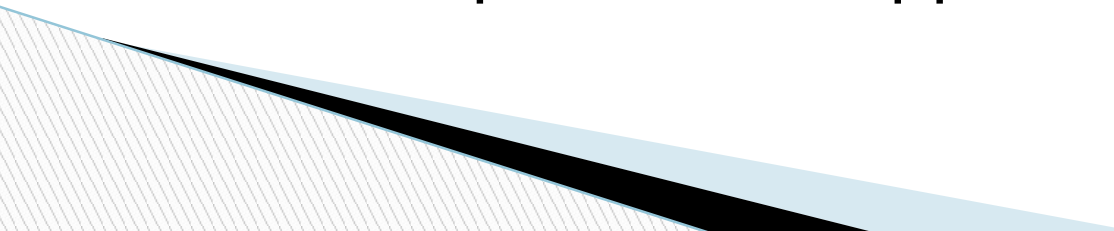
# Conclusion

- What have we learned ?
    - Need for multidisciplinary approach
  - New methods
    - Combine QUAL and QUAN approaches
    - Combine types of expertise
- 

# Cognitive interviewing

- Cognitive interviewing is an approach to evaluate sources of measurement error in survey questionnaires.

# General features of CI (Willis, 1999)

- a) Focuses mainly on the questionnaire, rather than on the entire survey administration process.
  - b) It explicitly focuses on the cognitive processes that respondents use to answer survey questions; herefore, covert processes that are normally hidden, as well as overt are studied.
  - c) For the conduct of the cognitive interview, volunteer subjects are recruited, and are interviewed either in a laboratory environment, or in some other private location.
  - d) Sampling procedures are different from those used in quantitative approach.
- 


# Model of CI

Tourangeau's (1984) model of CI:

## 1) COMPREHENSION OF THE QUESTION:

- a) Question intent: What does the respondent believe the question to be asking?
- b) Meaning of terms: What do specific words and phrases in the question mean to the respondent?

## 2) RETRIEVAL FROM MEMORY OF RELEVANT INFORMATION:

- a) Recallability of information: What types of information does the respondent need to recall in order to answer the question?
  - b) Recall strategy: What type of strategies are used to retrieve information? For example, does the respondent tend to count events by recalling each one individually, or does he/she use an estimation strategy?
- 

# Model of CI

## 3) DECISION PROCESSES:

- a) Motivation: Does the respondent devote sufficient mental effort to answer the question accurately and thoughtfully?
- b) Sensitivity/Social Desirability: Does the respondent want to tell the truth? Does he/she say something that makes him/her look "better"?

## 4) RESPONSE PROCESSES:

Mapping the response: Can the respondent match his or her internally generated answer to the response categories given by the survey question?

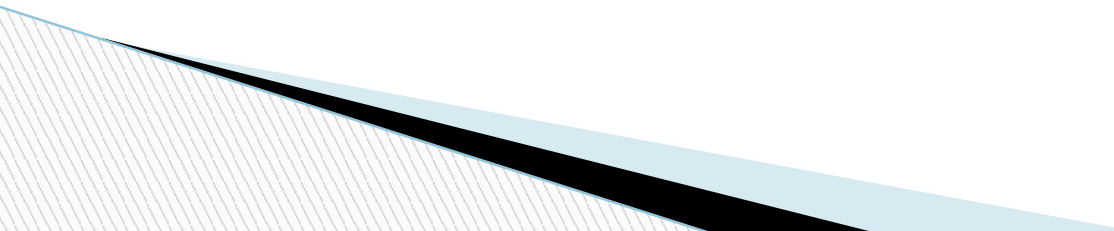
# CI techniques

- Think-aloud (Ericsson and Simon, 1980)
  - Cognitive/ Verbal probing (Willis, 1999)
- 



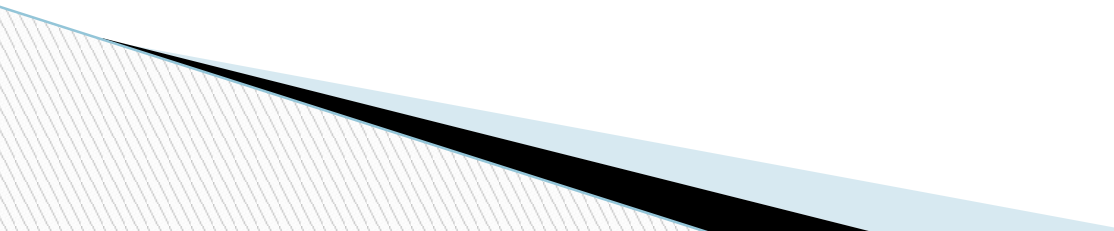
# CI techniques: Think-aloud

## **Advantages of the think-aloud technique:**

- a) Freedom from interviewer-imposed bias
  - b) Minimal interviewer training requirements
  - c) Open-ended format
- 

# CI techniques: Think-aloud

## Disadvantages of the think-aloud technique:

- a) Need for subject training
  - b) Subject resistance
  - c) Burden on subject
  - d) Tendency for the subject to stray from the task
  - e) Bias in subject information processing
- 

# CI techniques: Think-aloud

## Example

INTERVIEWER (reading survey question to be tested): How many times have you talked to a doctor in the last 12 months?

SUBJECT: I guess that depends on what you mean when you say “talked.” I talk to my neighbor, who is a doctor, but you probably don’t mean that. I go to my doctor about once a year, for a general check-up, so I would count that one. I’ve also probably been to some type of specialist a couple of more times in the past year - once to get a bad knee diagnosed, and I also saw an ENT about a chronic coughing thing, which I’m pretty sure was in the past year, although I wouldn’t swear to it. I’ve also talked to doctors several times when I brought my kids in to the pediatrician - I might assume that you don’t want that included, although I really can’t be sure. Also, I saw a chiropractor, but I don’t know if you’d consider that to be a doctor in the sense you mean. So, what I’m saying, overall, is that I guess I’m not sure what number to give you, mostly because I don’t know what you want.

# CI techniques: Verbal Probing

## **Advantages of the Verbal Probing technique:**

- a) Control of the interview
- b) Ease of training of the subject

## **Disadvantages of probing techniques:**

- a) Artificiality
  - b) Potential for Bias
- 

# CI techniques: Verbal Probing

## Example

*Comprehension/  
Interpretation probe:*

What does the term "outpatient" mean to you?

*Paraphrasing<sup>2</sup>:*

Can you repeat the question I just asked in your own words?

*Confidence judgment:*

How sure are you that your health insurance covers drug and alcohol treatment?

*Recall probe:*

How do you remember that you went to the doctor five times in the past 12 months?

*Specific probe:*

Why do you think that cancer is the most serious health problem?

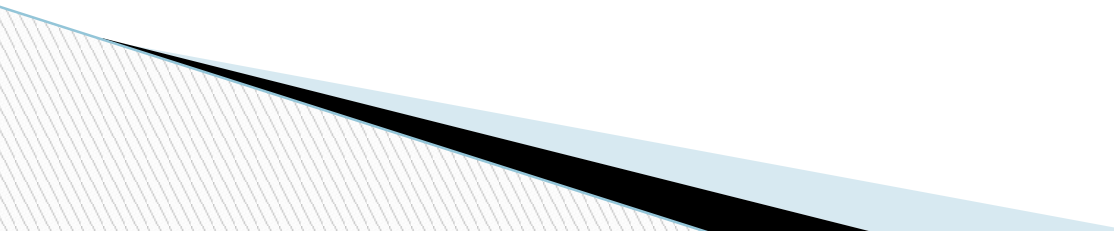
*General probes:*

How did you arrive at that answer?

Was that easy or hard to answer?

I noticed that you hesitated - tell me what you were thinking<sup>3</sup>

# CI techniques: Verbal Probing

- Concurrent probing: probing is done after every question asked
  - Retrospective probing: probing is done after entire questionnaire is administered
- 

# CI techniques: Verbal Probing

- a) Scripted probes: developed prior to the interview
  - b) Spontaneous probes: “thought up” probes during the interview
- 