Scale development for measurements in food consumer research (+ its place in the stepwise approach)

Training 2 Fruit consumption

Ivo A. van der Lans

Ohrid, June 22nd, 15:05-17:15

Schedule for this presentation

1. From idiosyncratic terms to common meaning
2. Exercise 2: Item generation
3. Selection of items: Data- versus expert-driven

Break

3. Selection of items: Data- versus expert-driven
4. Scale reduction
5. Cross-cultural research
6. Back to the stepwise approach

Cristina Mora

'**Qualitative paradigm** concentrates on investigating subjective data, in particular, the perceptions of people involved. The intention is to *illuminate* these perceptions and, thus, gain greater insight and knowledge'

'**Quantitative paradigm** concentrates on what can be measured. It involves collecting and analysing objective (often numerical) data that can be organised into statistics'

Generalisation by logic, reasoning, arguments

Generalisation by statistical inference

1. From idiosyncratic terms to common meaning

Qualitative

- Respondents express themselves using their own (idiosyncratic) terms

- Unstructured, semi-structured, interactive

- Researcher interprets *afterwards* and tries to capture the intended meaning of the respondents' terms

- Coding, categorization

Quantitative

- Respondents express themselves using the researcher's terms

- Structured and standardised

- Researcher has to make sure *in advance*, that respondents interpret his/her terms in his/her intended way

- Scale development

Grunert et al. (2007): Satisfaction with food-related life

One items (Disagree 1 – Agree 7)
- I am satisfied with my food-related life

Multiple items (Disagree 1 – Agree 7)
- My life is c...
- With r... of m...
- I am g...
- Food a... in daily life
- Food a... nts in my life
- When ...only disappointments
- I wish my meals were a much more plea...

Random error on individual statements is cancelled out
➔ **increases reliability**
➔ **increases validity**

Specificity of individual items is cancelled out
➔ **increases validity**

Finer grained scale
➔ **increase reliability**
➔ **increases validity**

LEI WAGENINGEN UR

FOCUS
FOOD CONSUMER SCIENCE IN THE BALKANS

STEP 1: Definition of the construct: what *is*, and what is *not* included

- Develop mini-theory:
  - based on thinking, discussing, literature, qualitative research
  - distinction from other, already existing, constructs
  - which aspects can be distinguished?

STEP 1

- **Satisfaction with food-related life**

- 'Quality of life' is not a clearly defined construct
  - objective indicators, and
  - subjective perceptions ➜ subjective well-being
    - affective component
    - cognitive component ➜ Satisfaction with Life Scale (global and domain-specific)

DEFINITION: 'overall evaluative assessment of that part of a person's life comprising, procurement, preparation and consumption of food and meals according to his/her criteria'

STEP 2: Generate pool of items

- based on thinking, discussing, literature, qualitative research, expert interviews
- how do others measure the same/similar construct(s)? (Handbook of Marketing Scales/Marketing Scales Handbook)
- redundancy is welcome!
  - include different shades of meaning
  - include positively and negatively phrased items
  - include multiple items for each and every aspect of the construct
- general rules for item construction
  - precise and clear, short, simple, no jargon, no double-barred questions, it is not bad to include no double negations
  - different response scales are not forbidden, but take care

- **Convenience orientation towards meal preparation**

- Candel (2001)
  - the degree to which a consumer is inclined to save time and energy as regards meal preparation'

- Scholderer & Grunert (2005)
  - additionally distinguish between physical and mental energy
  - (and also deal with convenience as regards decision-making, purchase, consumption, and cleaning up)

Candel (2001)

- The less physical energy I need to prepare a meal, the better
- The ideal meal can be prepared with little effort
- Preferably, I spend as little time as possible on meal preparation
- I want to spend as little time as possible cooking
- At home I preferably eat meals that can be prepared quickly
- It's a waste of time to spend a long time in the kitchen preparing a meal

Scholderer & Grunert (2005)

- **Attitude to convenience products**
  - We use a lot of ready-to-eat foods in our household
  - In our house, nibbling has taken over and replaced set eating hours
  - I use a lot of frozen foods in my cooking
  - I use a lot of mixes, for instance baking mixes and powder soup

- **Convenience product usage**
  - I use ready-prepared dishes that just need to be heated up
  - I snack instead of eating a big dinner
  - I eat fast food out

LEI
WAGENINGEN UR

- **STEP 3: Select reduced set of items based on qualitative judgments**

STEPs 4-8: Pilot study
- Churchill (1979) talks about pre-test

- **STEP 4: Administer items to sample of reasonable size (N=200, Grunert et al.: N=100)**

- STEP 5 Factor analysis
  - to check number of aspects/unidimensionality
  - to check which items seem to measure which aspects
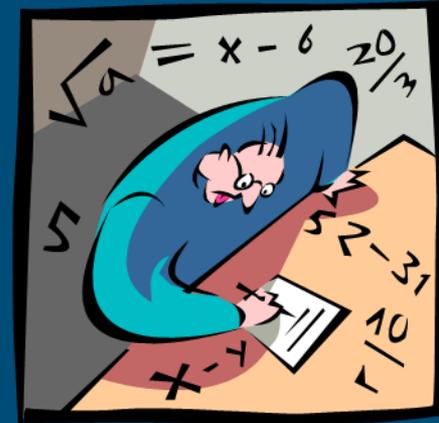
$$Y_j \approx \Sigma_r Z_r a_{jr} + U_j$$

Y_j items

Z_r aspects = *common* factors (capture the variance due to underlying (aspects of the) construct)

a_{jr} loadings of items on common factors (represent contribution of common factors to scores on item)

U_j *unique* factors (capture *measurement error* = variance that is unique to the item)

  - delete (or reformulate) items that load on *more than one* or *no* factor

- **STEP 6   Reliability analysis**

  $$X_{observed} = X_{true} + (X_{systematic} + X_{random})$$

  - X sum or mean score across items that load high on the same common factor (*reversed* scores on negatively phrased items)

  - $reliability = r^2(X_{observed}, X_{true} + X_{systematic})$

    - $X_{random} = 0$ $\Rightarrow$ $X_{observed}$ is perfectly *reliable*

    - $X_{systematic} + X_{random} = 0$ $\Rightarrow$ $X_{observed}$ is perfectly *valid*

LEI WAGENINGEN**UR**

FOCUS
FOOD CONSUMER SCIENCE IN THE BALKANS

- Reliability estimate: Cronbach's α (internal consistency reliability)

$$\alpha = \frac{(k)}{(k-1)} \times \left(1 - \frac{\sum_{i=1}^{k} \text{var}(Y_i)}{\sum_{i=1}^{k} \text{var}(Y_i) + 2\sum_{i>j} \text{cov}(Y_i Y_j)}\right)$$

- all items uncorrelated ➜ $\alpha = 0$
- all items perfectly correlated and equal variance ➜ $\alpha = 1$

- Test-retest reliability: correlation between scores on *same scale* at two different points in time

- *true* scores may change in between (deflates reliability estimate)
- respondents may remember their answers on the first occasion (inflates reliability estimate)
- estimate depends on time between two administrations
- not possible for scales that should measure first reaction

- **Rules of thumb for reliability**
  - early stages $\Rightarrow$ 0.50 to 0.60 suffices
  - basic research $\Rightarrow$ 0.80 suffices
  - applied research $\Rightarrow$ 0.90 – 0.95 should be attained
  - reliability depends on population! (in which population do we want to make reliable distinctions?)
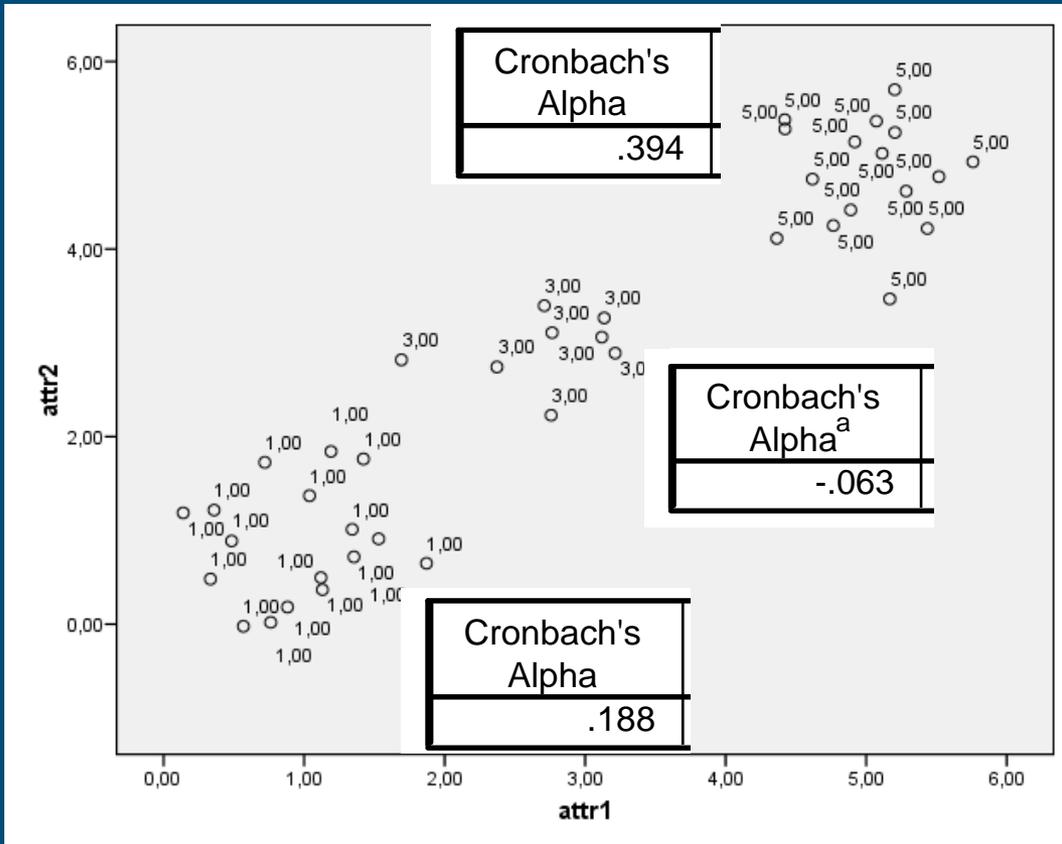
- **STEP 7 Adjust scale if it is not reliable enough**
  - delete items to increase reliability (iterative process)
  - *never, ever*, delete an item without looking at its content
  - reformulate items
  - think about new items
  - do not increase the reliability at the expense of validity!

■ Reliability for distinguishing between what?

- **STEP 8 Check validity**
  - Face/content validity
    does the set of items look OK?
  - Criterion validity
    does the scale predict what it is supposed to predict?
    - concurrent versus predictive criterion validity
  - Construct validity
    does the scale really measure the construct?
    - Convergent validity
      high correlations with different scales for the same construct
    - Discriminant validity
      not too high correlations with (similar) scales for other constructs
    - Nomological validity
      correlate with scales for other constructs as expected

- **STEP 9   Larger/main study**
  - unidimensionality
  - reliability
  - validity

- **STEP 10   Develop norms**
  - who has a low fruit consumption?
  - who is highly satisfied with his/her food-related life?

- Data analysis in STEPs 4-9 has become more sophisticated by the introduction of *confirmatory factor analysis* in marketing and consumer research (e.g., Steenkamp and Van Trijp, 1991)

3. Selection of items: Data- versus expert-driven

- **Critique on data-driven approach**

  - disproportional emphasis on reliability
  - disproportional emphasis on criterion and construct validity
  - disproportional emphasis on multi-item scales

  - all very much enhance by adopting confirmatory factor analysis as 'golden standard'

- **Rossiter's (2002) C-OAR-SE procedure tries to remedy this**

  - more emphasis on expert judgment
  - more emphasis on qualitative research in pretest
  - more emphasis on construct definition
  - more emphasis on content validity

■ **C-OAR-SE**

- Construct *definition* in terms of Object, Attribute, and Rater Entity

- Object classification …
  - Concrete singular (e.g. banana)

  - Abstract collective (e.g., fresh fruit), has *constituents*

 - Abstract formed (e.g., food-related life), has *components*

  … by means of e.g. open interviews, or focus groups, and expert discussions

■ C-OAR-SE

- ● Attribute classification

  - concrete singular
    (e.g., buying intention)(one item)

  - eliciting (e.g., convience orientation)
    (reflective items, overall reliability)

  - formed (e.g., quality perception)
    (formative components, reliability per component)
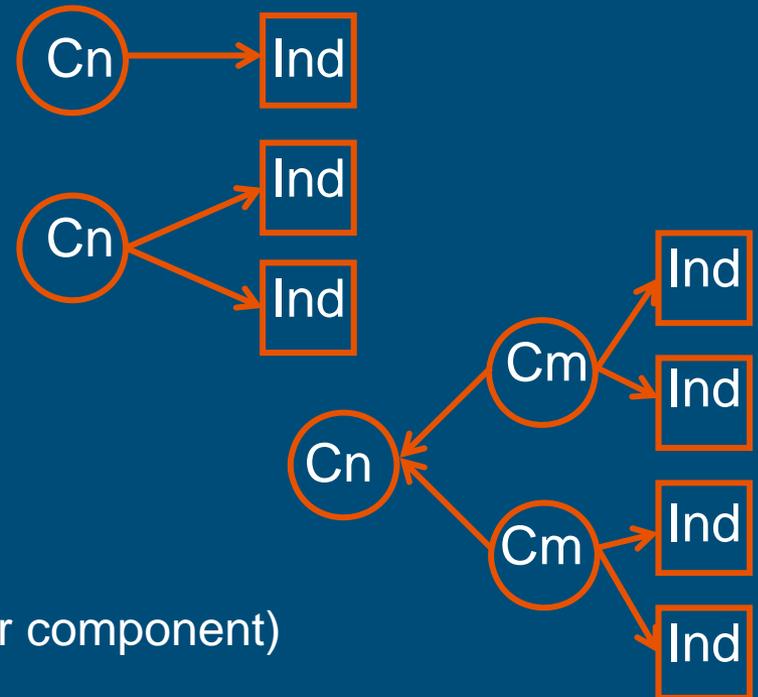
3. Selection of items: Data- versus expert-driven

- Rater identification
  - individual consumers
  - expert judges
  - sample of consumers

  … also part of the construct definition!

- Scale formation
  - Combine objects and attributes into items
    - similar to idea of mapping sentences in facet theory
  - Select response scale
  - Pre-test for comprehension
  - Pilot for unidimensionality of eliciting-attribute items
  - Check reliability (which crucially depends on rater entity)
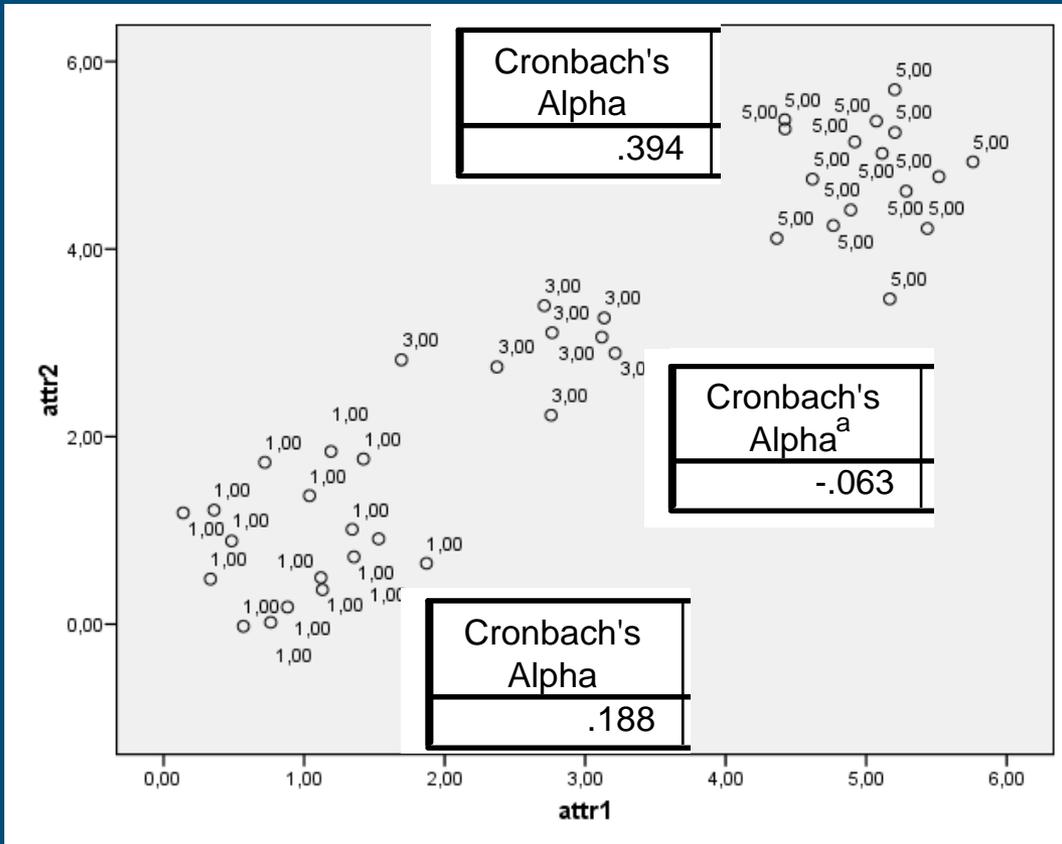
- Enumeration
  - Calculate index or average …. or
  - Transform to meaningful range (0 to 10, -5 to 5)

■ Reliability for distinguishing between what?

Reliability Statistics

| Cronbach's Alpha	N of Items
.965	2

4. Scale reduction

- **Satisfaction with food-related life**
  - not more than three to eight items (self-administered, telephone-based, face-to-face)
  - how about other scales with abstract collective/formed objects and eliciting/formed attributes?

- **Scale-reduction practices mainly focus on maintaining a high internal consistency (high reliability)**
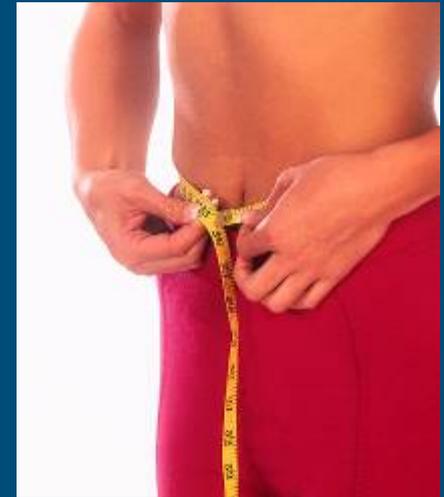
4. Scale reduction

- Stanton et al. (2002) propose three key aspects of item quality
  - Internal item qualities
    - distribution (discrimination and threshold)
    - convergent validity with full-length scale
  - External item qualities
    - nomological and discriminant validity
  - Judgmental item qualities
    - comprehensibility
    - semantic redundancy
    - invasiveness

5. Cross-cultural research

- **What can go wrong?**
  - Das Leben in vollen Zügen geniessen (German)

- **Interaction needed between experts on:**
  - study
  - questionnaire design
  - linguistics
  - culture

5. Cross-cultural research

- **Translation**
  - backtranslation (bilingual native speaker source) (after translation by bilingual native speaker target)
  - parallel/double translation

- **Pretest**
  - ask for comprehension
  - ask for meaning

- **Documentation**
  - especially when study will be repeated

5. Cross-cultural research

- **Category equivalence**
  - status of jobs, marital status, soft drinks

- **Functional equivalence**
  - fruit consumption may have different functions in different cultures
    - snack or dessert
    - social activity or something like taking medicine

- **Conceptual equivalence**
  - different connotations
  - reflect different intensities

- **Together they form construct equivalence**
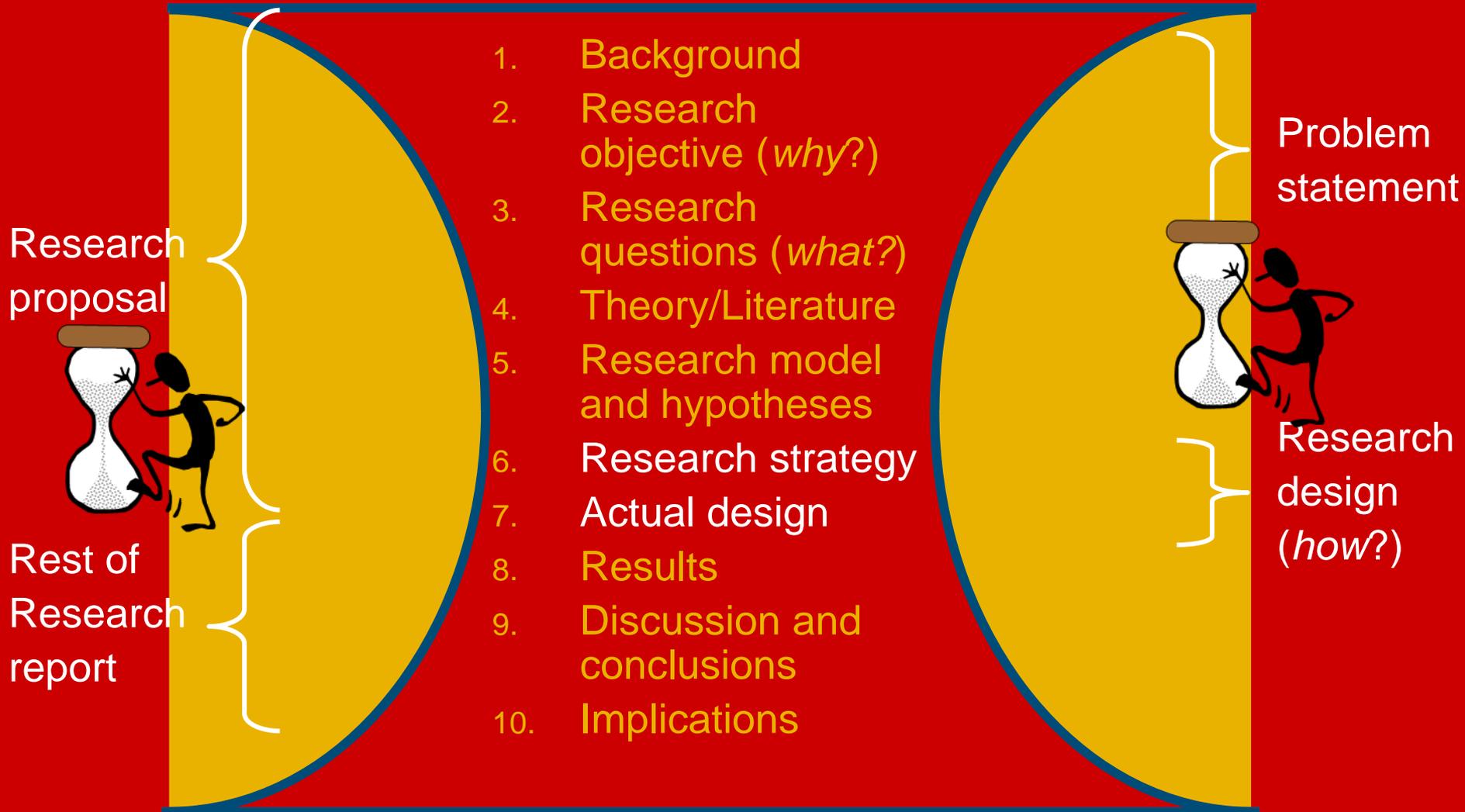 - equivalence is important for both the item and the response scale

5. Cross-cultural research

- Multi-item scales allow for checks on the culture-specificity of particular items

- (Multi-group) confirmatory factor analysis provides a tool to do this in a structured way, … and to look at cultural differences in construct validity

- But still, …

6. Back to stepwise approach

1. Background
2. Research objective (*why*?)
3. Research questions (*what?*)
4. Theory/Literature
5. Research model and hypotheses
6. Research strategy
7. Actual design
8. Results
9. Discussion and conclusions
10. Implications

Research proposal

Rest of Research report

Problem statement

Research design (*how*?)

- **Standardized questions and multi-item scales**
  - fit in quantitative *research strategies*: Surveys and experiments
  - deal with the *actual design*: *how to measure* the (theoretical) constructs in the *research model*
  - enable us to assess univariate statistics for, and associations between, constructs in the *research model* and to test *hypotheses*

- **Once data-collection has started … there is no way back!**
  - pre-test and pilot study are indispensable

6. Back to stepwise approach

- **Grunert et al. (2007): Satisfaction with food-related life**

- *Background*: in which situations and in which theories does the construct play a role
  - "Food is an essential component of quality of life; an unacceptable or unpalatable diet can lead to poor food and fluid intake, resulting in weight loss and undernutrition and a spiral of negative health effects" (American Dietetic Association, 2005)
  - how do objective indicators influence satisfaction with food-related life?

- *Objective*: develop a measure for satisfaction with food-related life

- *Research questions*
  - what is satisfaction with food-related life?
  - which aspects does food-related life consist of?
  - which items measure food-related life in a reliable and valid way?

Some references

- **Marketing-research textbooks**
  - Aaker, D., Kumar, V. & Day, G. (2001). *Marketing Research* (7th Ed.). New York: Wiley.
  - Churchill, G.A., Jr. & Iacobucci, D. (2002). *Marketing Research: Methodological Foundations* (8th Ed.). Fort Worth: Harcourt.
  - Malhotra, N.K. & Birks, D.F. (2007). *Marketing Research: An Applied Approach* (3rd European Ed.). Harlow, UK: Prentice Hall.

- **Scale development**
  - Bearden, W.O. & Netemeyer, R.G. (1999). *Handbook of Marketing Scales: Multi-item Measures for Marketing and Consumer Behavior Research* (2nd Ed.). Thousand Oaks: Sage.
  - Bruner, G.C. (1992-2009). *Marketing Scales Handbook* (Vol. I-V). AMA/Thomson.
  - Candel, M.J.J.M. (2001). Consumers' convenience orientation towards meal preparation: Conceptualization and measurement. *Appetite, 36* (1), 15-28.
  - Churchill, G.A., Jr. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research, 16*, 64-73.
  - Rossiter, J.R. (2002). The C-OAR-SE procedure for scale development in marketing. *International Journal of Research in Marketing, 19*, 305-335.

Some references

- Scholderer, J. & Grunert, K.G. (2005). Consumers, food and convenience: The long way from resource constraints to actual consumption patterns. *Journal of Economic Psychology, 26*, 105-128.
- Stanton, J.M., Sinar, E.F., Balzer, W.K. & Smith, P.C. (2002). Issues and strategies for reducing the length of self-report scales. *Personnel Psychology, 55*, 167-194.
- Weijters, B., Geuens, M. & Schillewaert, N. (2009). The proximity effect: The role of inter-item distance on reverse-item bias. *International Journal of Research in Marketing, 26*, 2-12.

Cross-cultural
- Douglas, S.P. & Craig, C.S. (2007). Collaborative and iterative translation: An alternative approach to back translation. *Journal of International Marketing, 15* (1), 30-43.
- Steenkamp, J.-B.E.M. & Ter Hofstede, F. (2002). International market segmentation: Issues and perspectives. *International Journal of Research in Marketing, 19*, 185-213.

Confirmatory factor analysis
- Steenkamp, J.-B.E.M. & Van Trijp, H.C.M. (1991). The use of LISREL in validating marketing constructs. *International Journal of Research in Marketing, 8*, 283-299.

Cross-cultural & Confirmatory factor analysis
- Steenkamp, J.-B.E.M. & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research, 25*, 78-90.

Questions?

© Wageningen UR