2007-06-26

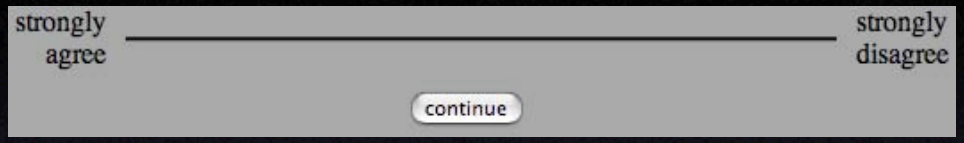
University of Zurich

Improving Data Quality in Web Surveys with Visual Analogue Scales

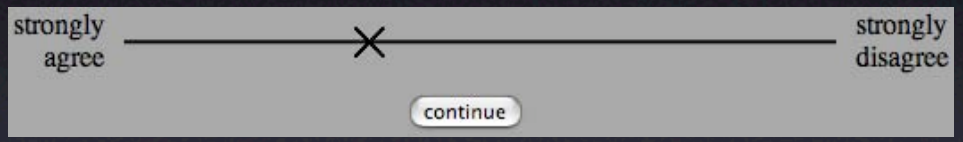
Frederik Funke & Ulf-Dietrich Reips (University of Zurich, Switzerland)

Visual Analogue Scales

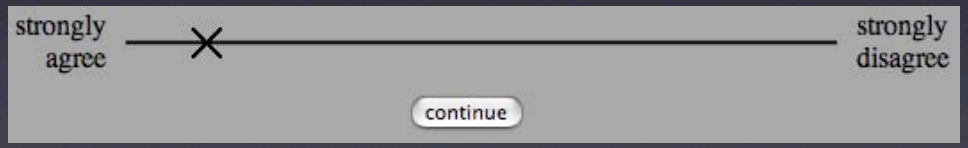
- on load:

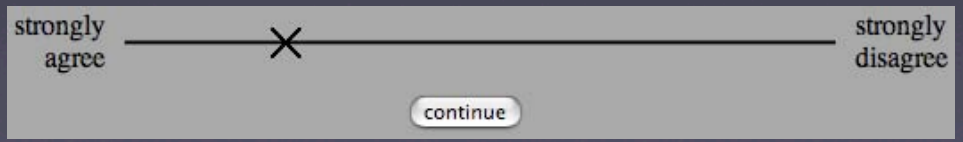


- indicator appears after first click:



- can be changed:





Overview

- Web experiment 1:
  data quality obtained with VAS

- Web experiment 2:
  re-test reliability

- Web experiment 3:
  how to compare VAS to categorical scales

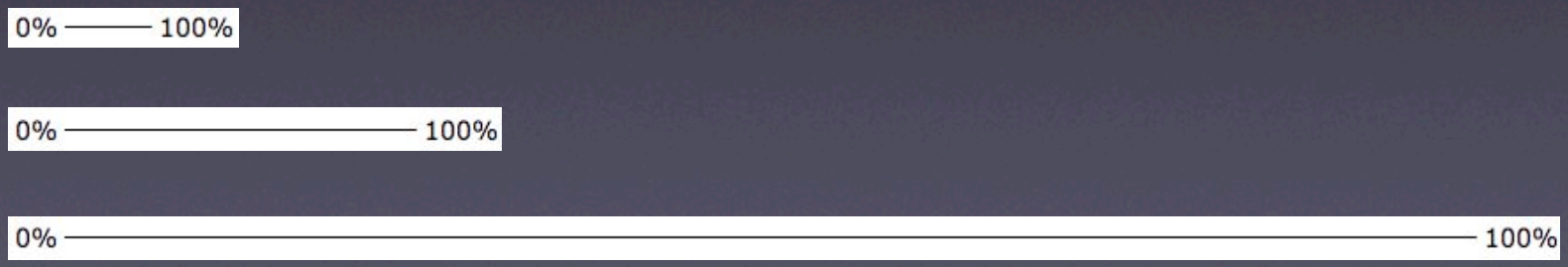
Experiment 1: Test for Equidistance

- How can data from VAS be interpreted?
- Do VAS produce data on interval level?
- What statistical procedures can be applied to data collected with VAS?

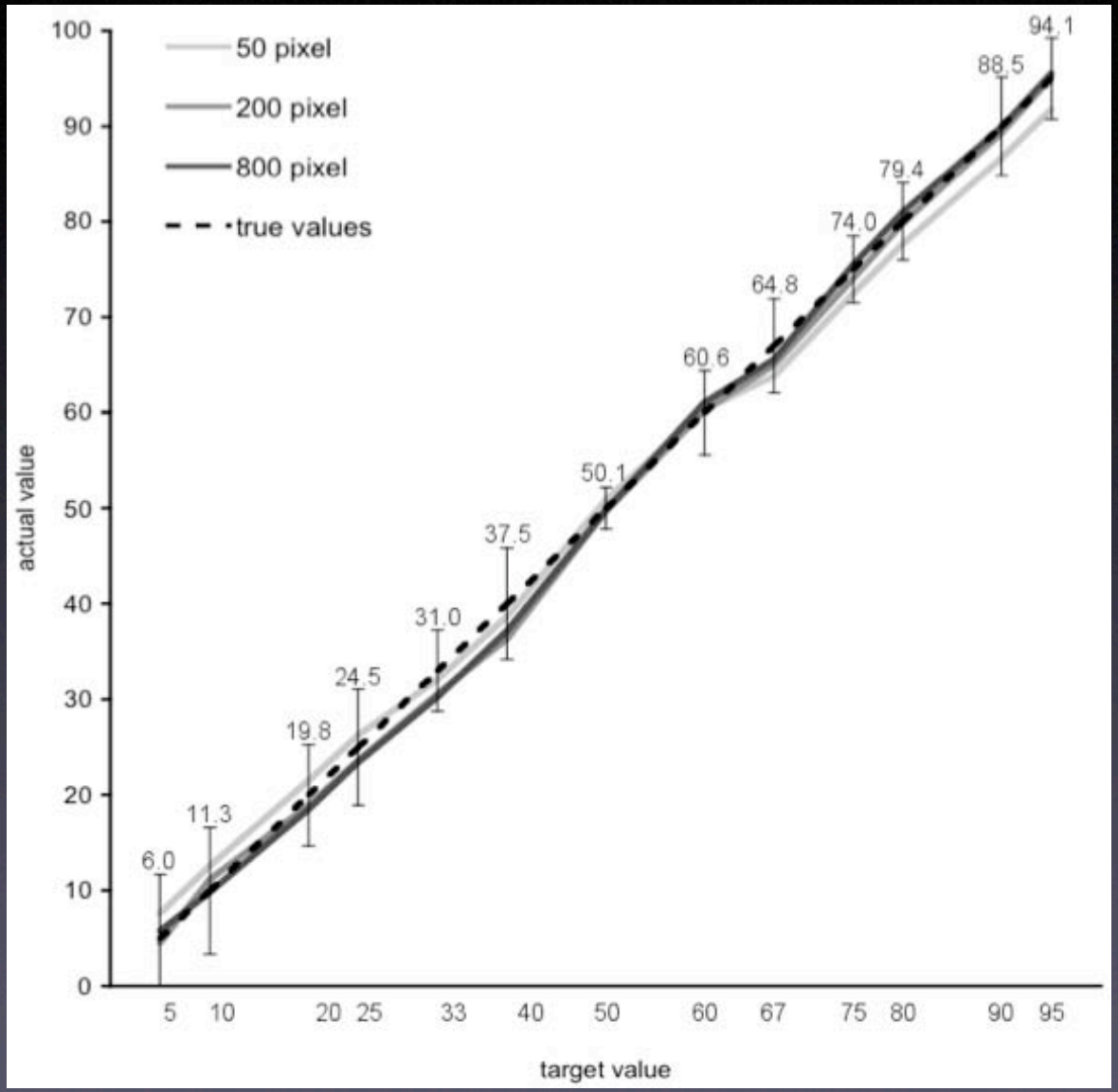
- interval level: difference between data points are meaningful; they do not only give information about a hierarchical order, but about the intensity of a difference; equal changes in intensity correspond to equal changes in data

Experiment 1: Test for Equidistance

- self-selected student sample
  between-subjects design
  n = 355
- 13 different target values (presented twice in reverse orders) had to be located on the VAS: 50%, 75%, 10%, 33%, 80%, 95%, 25%, 67%, 40%, 5%, 60%, 90%, 20%
- 3 VAS lengths: 50 pixel, 200 pixel & 800 pixel



Experiment 1: Test for Equidistance



Experiment 1: Test for Equidistance

- test for equidistance
  - to be equidistant: difference between 2 target values should correspond to difference between 2 actual values, regardless of position

interval	width		
	equidistant	realized	Δ
5% - 10%	5.0	*M*=5.0 (*SD*=8.3)	0.0
20% - 25%	5.0	*M*=4.4 (*SD*=5.5)	-0.6
75% - 80%	5.0	*M*=5.3 (*SD*=5.1)	0.3
90% - 95%	5.0	*M*=5.2 (*SD*=7.5)	0.2

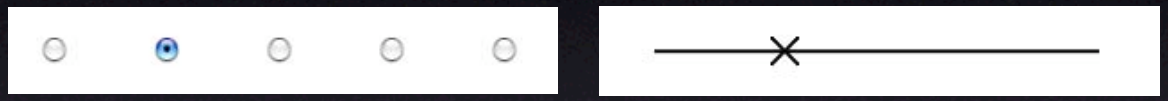
Experiment 1: Test for Equidistance

- conclusions
  - equal numerical intervals approximate equal segments on the VAS
    - strong evidence that data collected with VAS are equidistant, on the level of an interval scale
    - statistical procedures that require data on interval level can safely be applied
  - as length has no great effect, VAS should be robust to difference in appearance due to different screen sizes

Experiment 2: Re-Test Reliability

- Is measurement error different for VAS than for categorical scales?
- Do ratings with VAS produce more random measurement error?
- Do VAS' fine distinctions overtax respondents, resulting in greater variance of data?

Experiment 2: Re-Test Reliability

- 40 item inventory on personality
- student sample; n = 60
- 
- difference between t(1) and t(2) to determine re-test reliability

Experiment 2: Re-Test Reliability

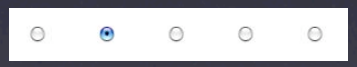
- with both scales, data running from 0 to 100
- mean difference between 2 measurements

   *M* = -0.8 (*SD* = 3.5)

   *M* = -0.9 (*SD* = 2.9)

  F (1, 59), *ns*

- re-test reliability (mean for all 40 items)

   Alpha = .80

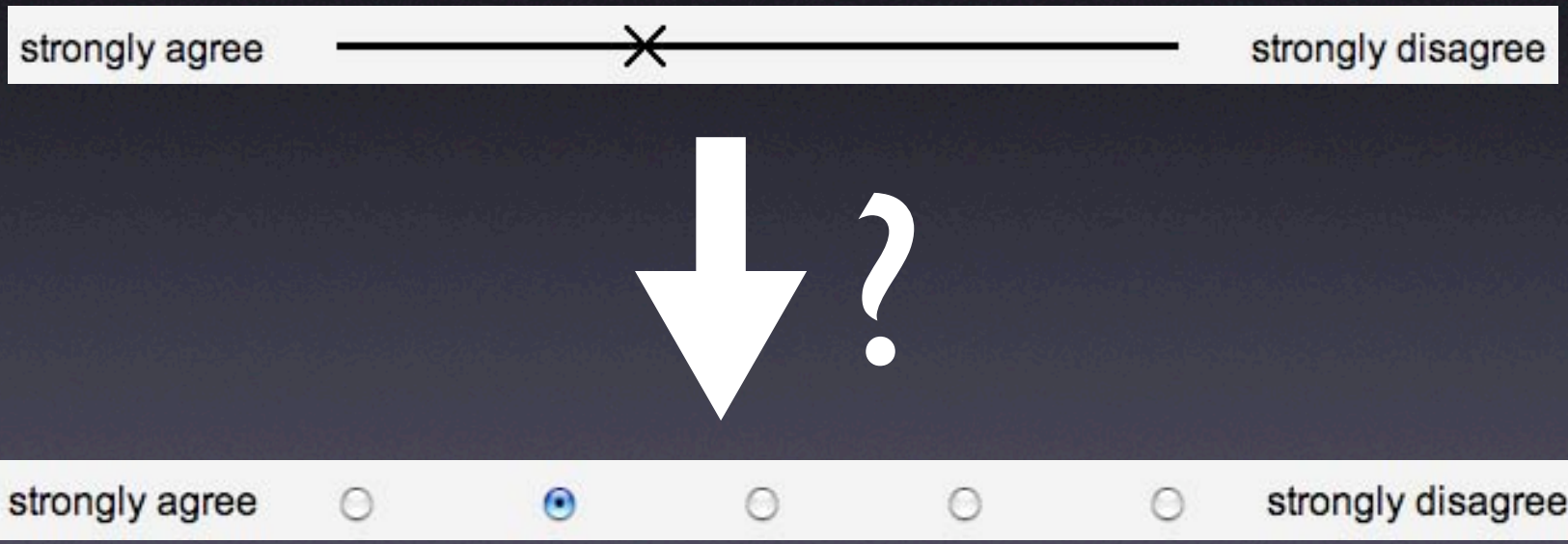
   Alpha = .85

Experiment 2: Re-Test Reliability

- conclusion
  - re-test reliability is slightly higher with VAS
  - variance with VAS is lower than with 5-point-categorical scale
  - VAS is used in a consistent way
  - respondents are not overtaxed with making a judgement on the VAS

Experiment 3: Comparing VAS to Categorical Scales

- How can data from VAS be compared to existing data from categorical scales?
- What is the appropriate categorization?



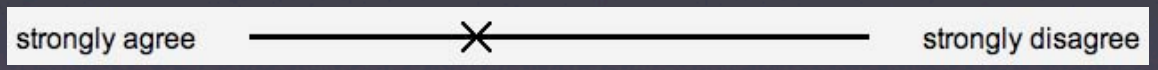
Experiment 3: Comparing VAS to Categorical Scales

- sample: 576 participants of the 33rd congress of the German Sociological Association

- 5-pt 

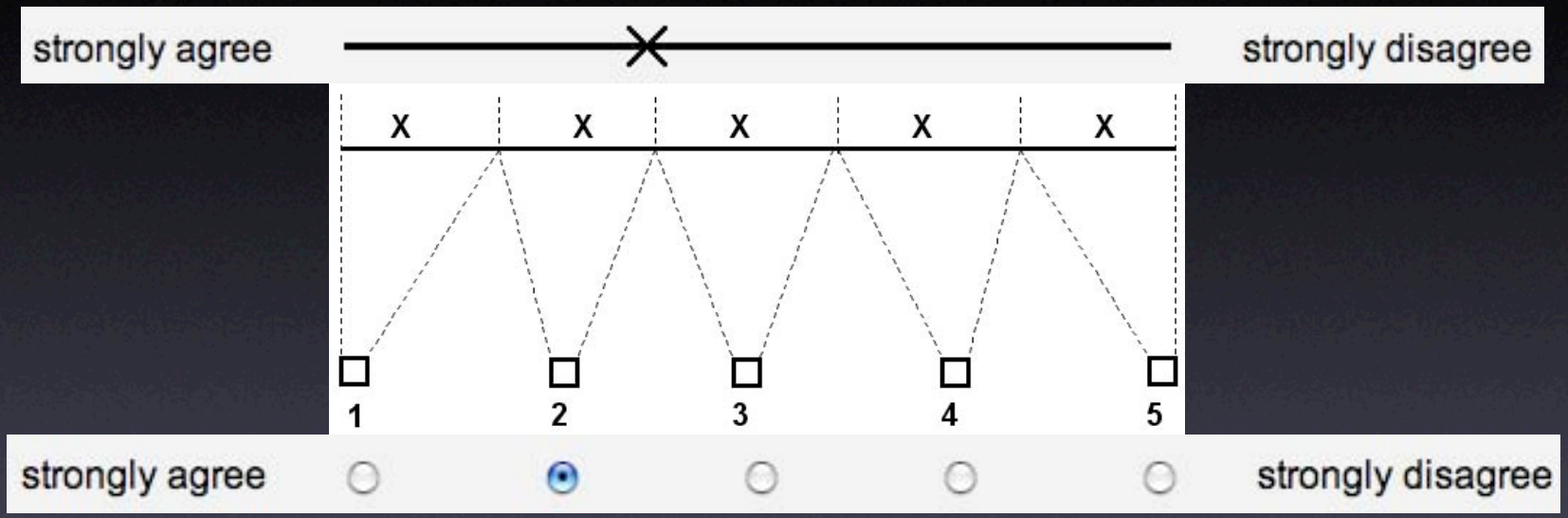
- 7-pt 

- 9-pt 

- VAS 

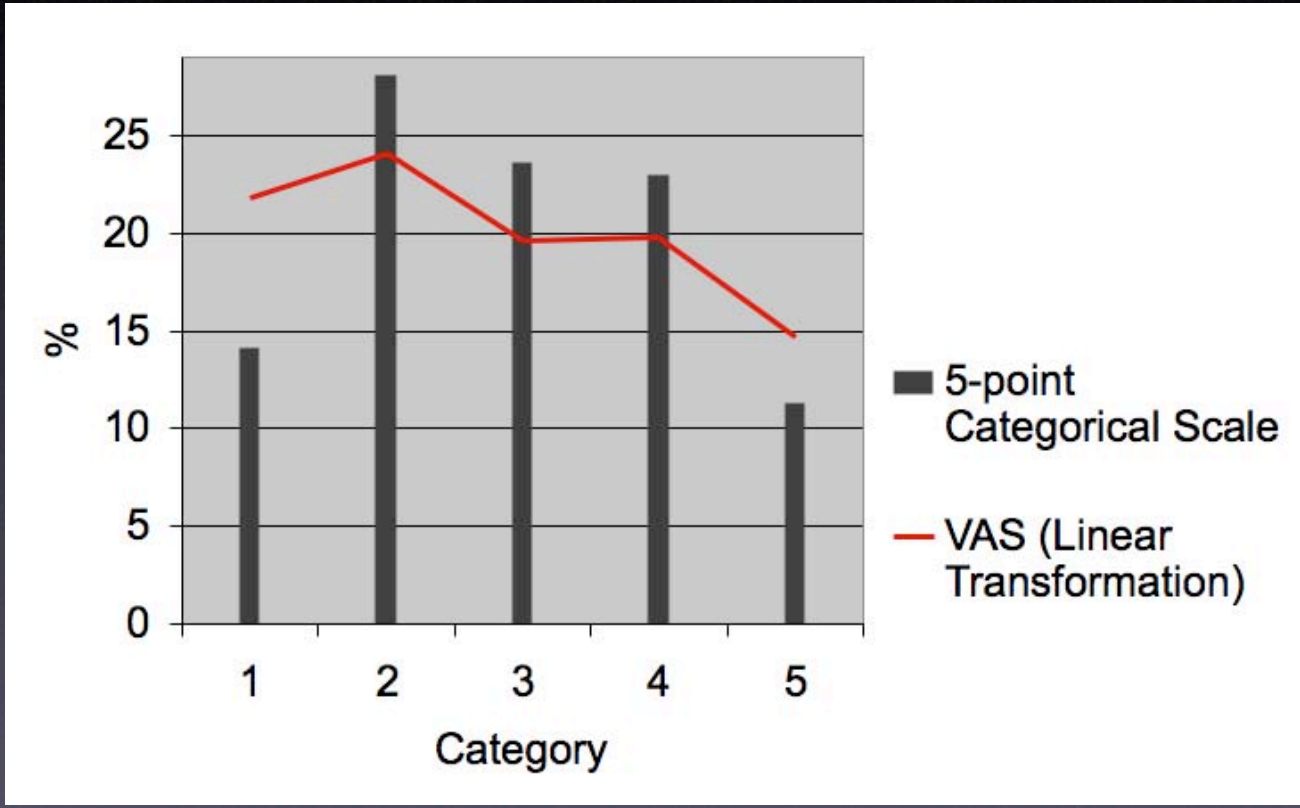
Experiment 3: Comparing VAS to Categorical Scales

- linear categorization of VAS



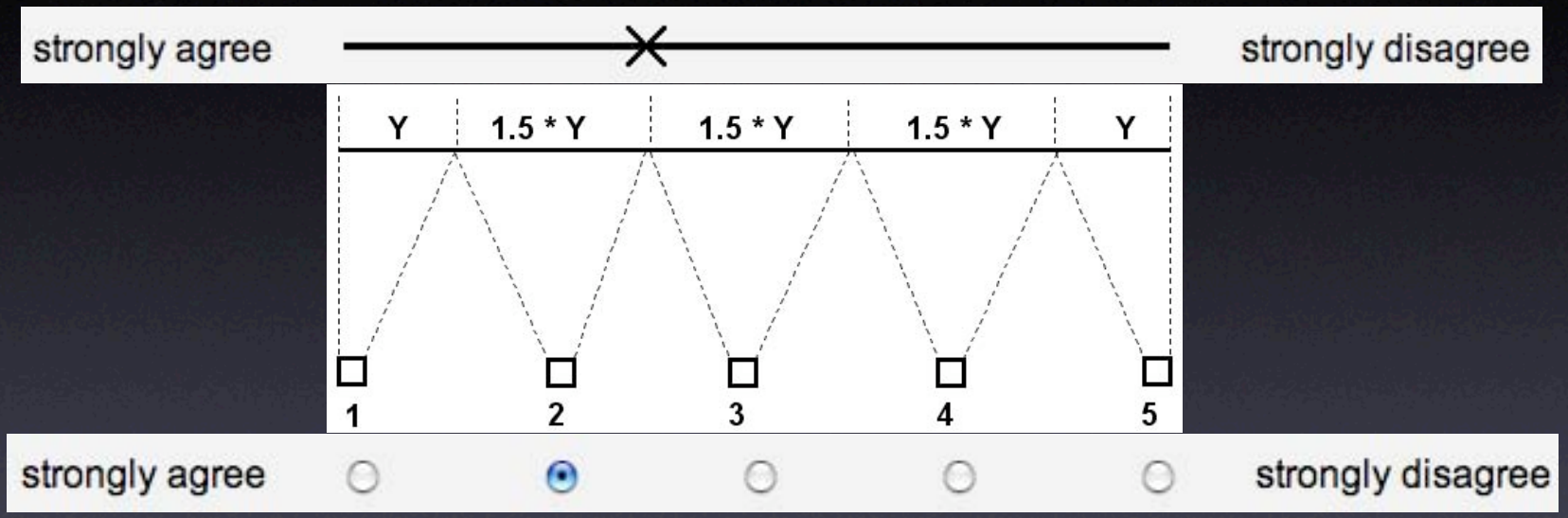
Experiment 3: Comparing VAS to Categorical Scales

- linear categorization of VAS



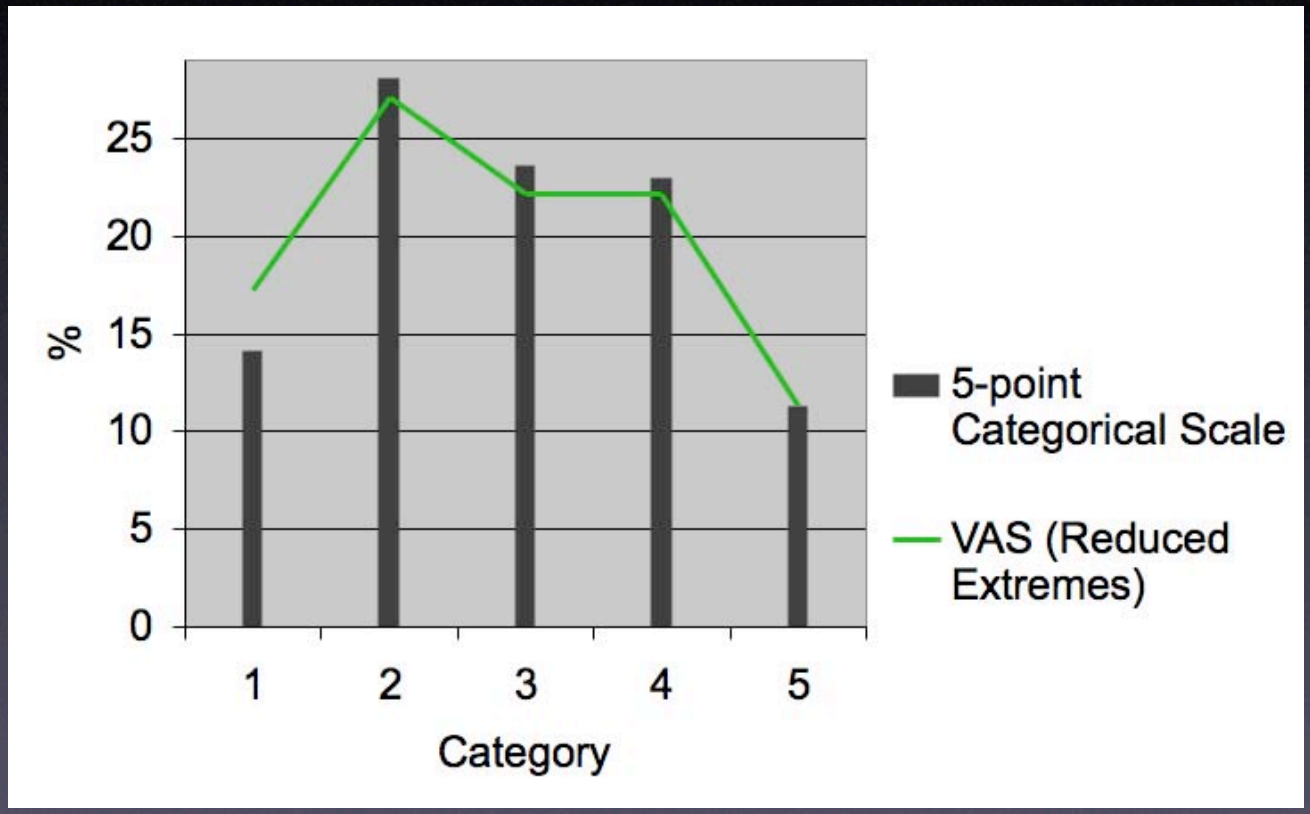
Experiment 3: Comparing VAS to Categorical Scales

- categorization of VAS with reduced extremes



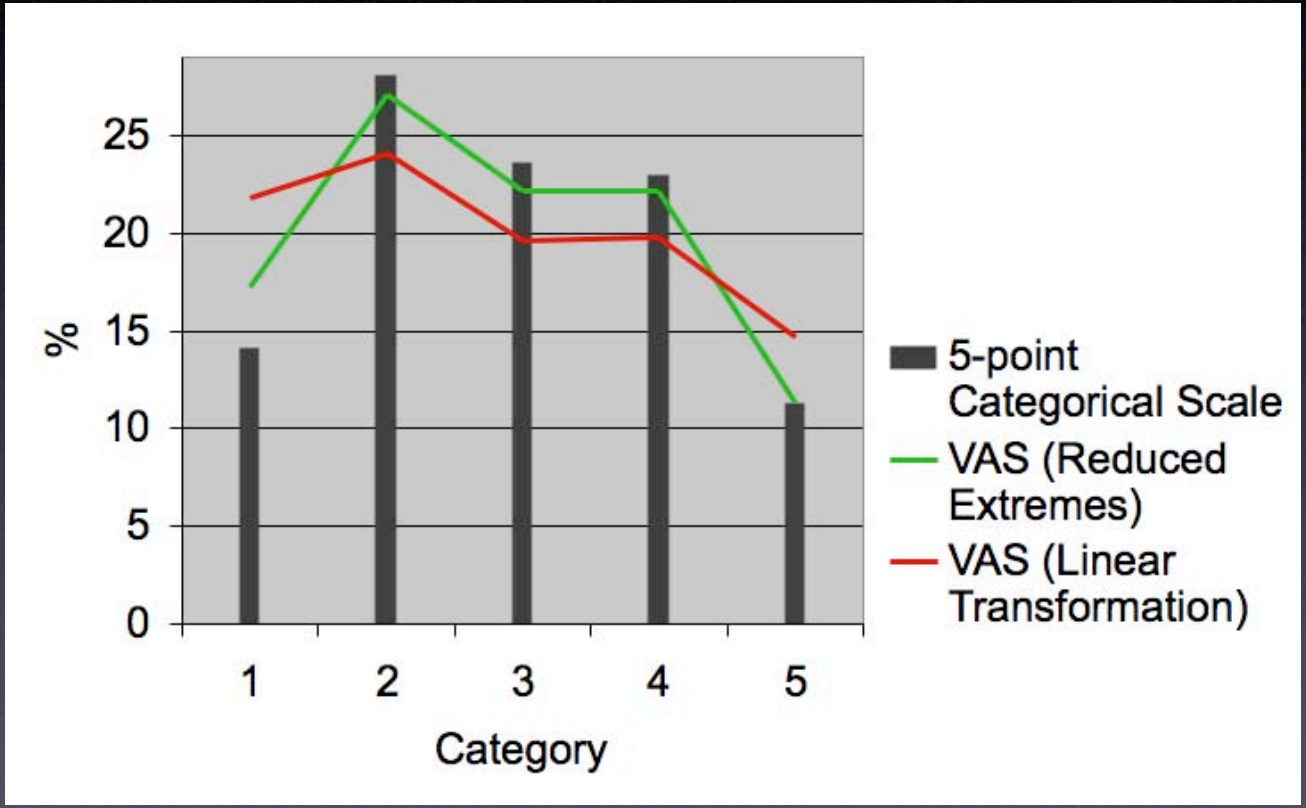
Experiment 3: Comparing VAS to Categorical Scales

- categorization of VAS with reduced extremes



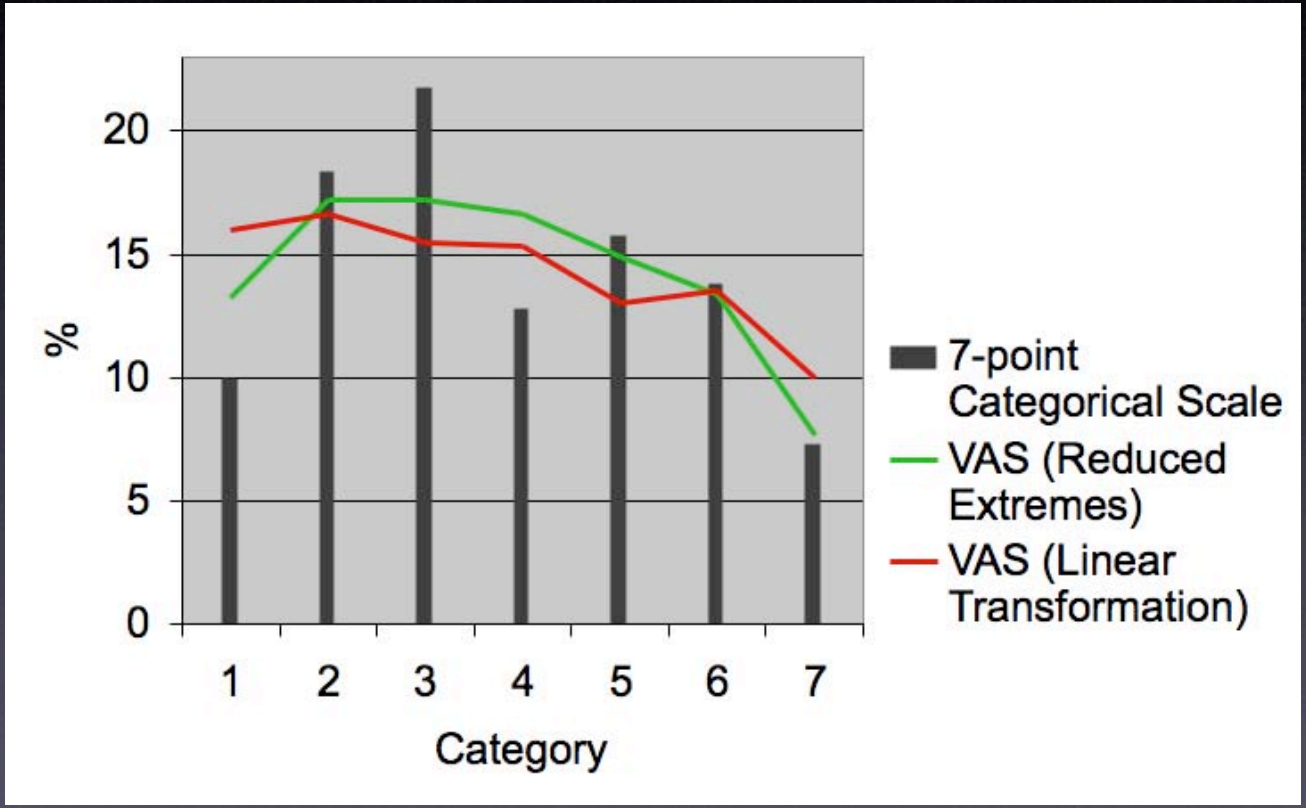
Experiment 3: Comparing VAS to Categorical Scales

- comparison of categorizations into 5 categories



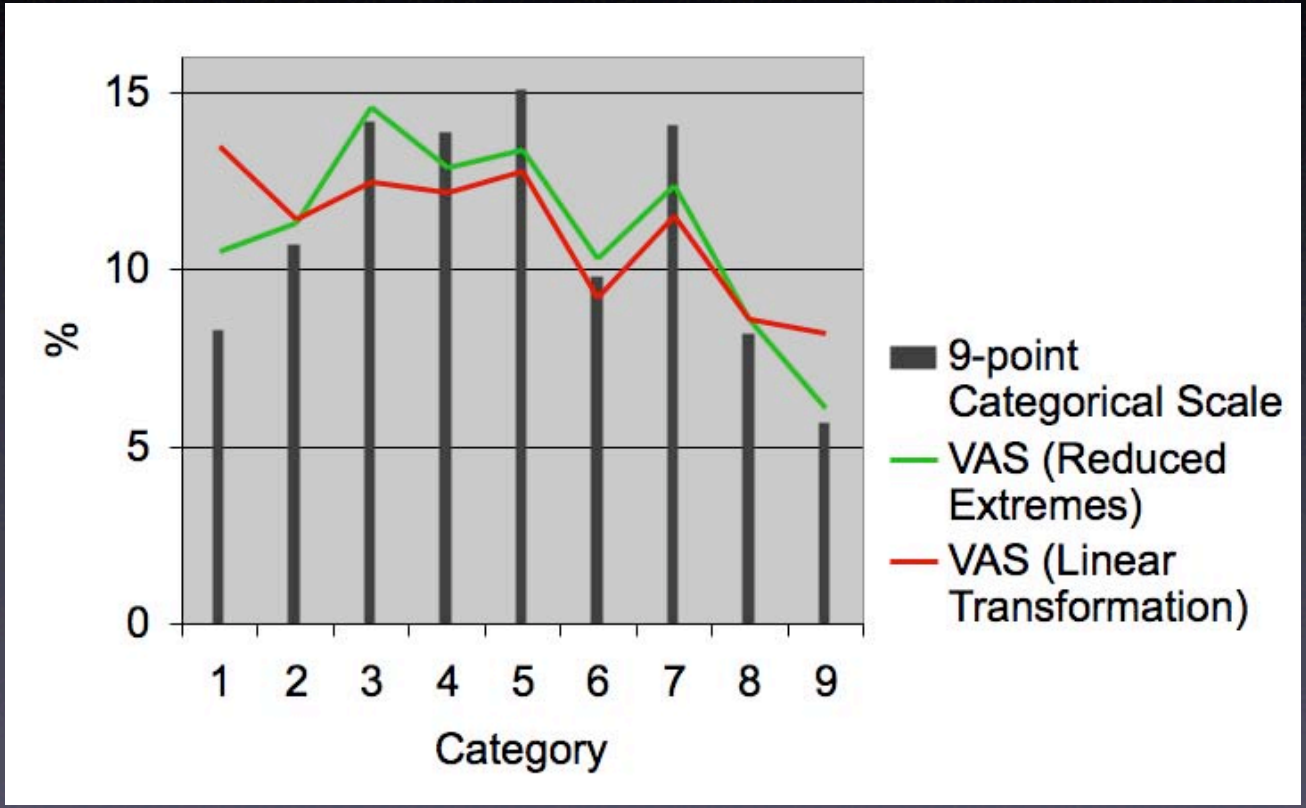
Experiment 3: Comparing VAS to Categorical Scales

- comparison of categorizations into 7 categories



Experiment 3: Comparing VAS to Categorical Scales

- comparison of categorizations into 9 categories



Experiment 3: Comparing VAS to Categorical Scales

- findings experiment 3
  - transformation with reduced extremes leads to greater accordance between VAS and categorical scales than linear transformation
  - difference between measurement with VAS and categorical scales is systematic

Combination of Experiment 1 & 3

- with data from VAS reaching interval level, they can be used as a benchmark to describe the width of categorical scales' categories
  - width of categories can be better described with VAS after categorization with reduced extremes
  - as respondents avoid picking extreme categories with categorical scales, extreme categories are systematically underrepresented
  - strong evidence that categorical scales produce ordinal data only

Conclusion

- data from VAS approximate an interval scale level
- rating with VAS is reliable
- measurement with categorical scales systematically differs from measurement with VAS
- when frequencies from VAS are to be compared to frequencies from categorical scales, transformation with reduced extremes is appropriate
- VAS should be applied for measurement of continuous variables

VAS Generator

- free Web service: http://www.vasgenerator.net