F. Funke[a] & U.-D. Reips[b] (University of Zurich)

Visual Analogue Scales in Online Surveys: Non-Linear Data Categorization by Transformation with Reduced Extremes

Introduction

Visual Analogue Scales (VAS) are graphic rating scales. They were first described by Hayes and Patterson in 1921. In most cases, a VAS is a simple horizontal line with verbal anchors on each end. Respondends convey their attitude or level of accordance by marking the point on the line they think is most appropriate. These scales have proven a highly reliable and valid instrument in surveys (Flynn et al., 2004).
Online surveys lend themseves to the use of VAS (see Figure 1). In contrast to paper and pencil interviews – where reading out data takes up a lot of time and is prone to errors – the readout in online surveys occurs automatically. The only client-side requirement is that the technology used to create the scales (e.g. JavaScript, Flash or Java) has to be enabled in the user's web browser.

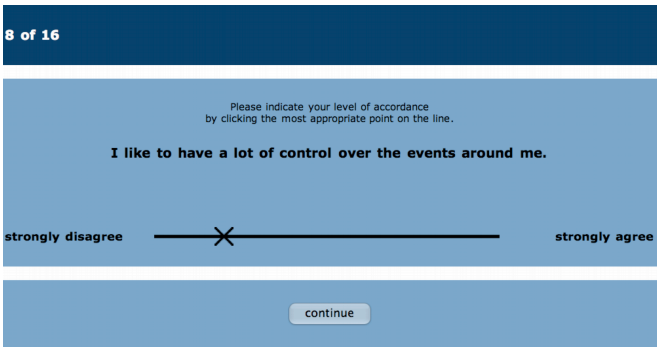

Figure 1: Visual Analogue Scale in an online survey

From a statistical point of view, the main advantage is that VAS generate data on interval scale level and therefore meet important requirements for the applicability of parametric procedures (Nyrén et al., 1987).
To understand the influence of VAS on the respondent's way of answering questions we conducted two experiments.

Experiment 1 – Relation Between Categorical Scales and VAS

In our first web experiment 667 participants were randomized to three different conditions to rate 16 items on power motivation (Schmidt & Frieze, 1997). The only difference between the experimental conditions consisted in the applied rating scales: Either a 4-point categorical scale, a 8-point categorical scale or a VAS.
To be able to compare the frequencies of categorical scales with data from VAS, VAS values need to be categorized. The commonly used way of categorizing VAS is linear transformation. Here, equal intervals form one category. Figure 2 shows linear transformation from VAS to a 4-point scale.

[a] email@frederikfunke.de
[b] u.reips@psychologie.unizh.ch

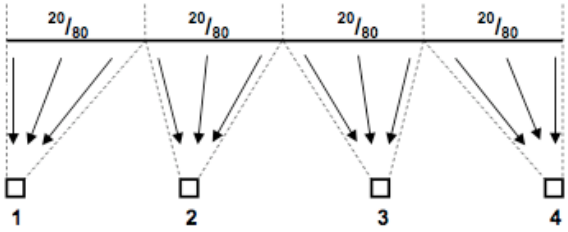

Figure 2: Linear transformation of VAS to 4-point scale

When comparing the frequencies between categorized VAS and the original categorical scale, we found typical systematic differences in the distribution, especially concerning the extreme categories (Figure 3).

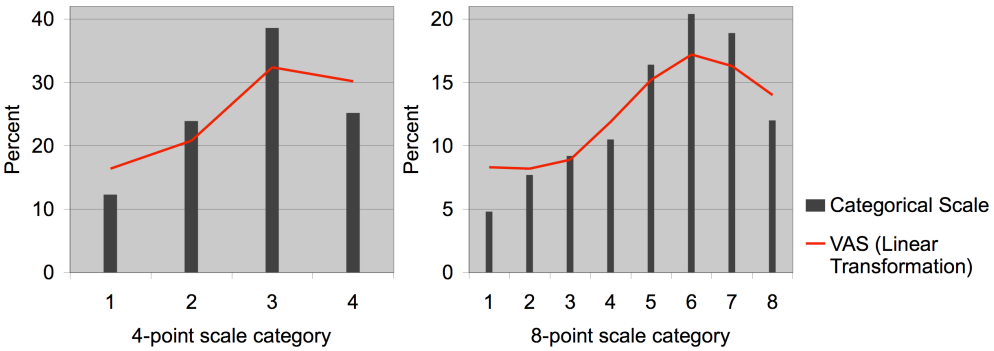

Figure 3: Frequencies of categorical scales and VAS after linear transformation

To gain greater accordance, we categorized the VAS in a different way: „Transformation with reduced extremes“ (Figure 4).

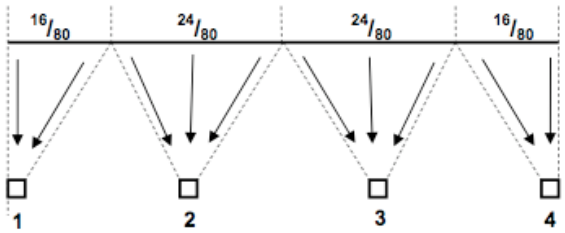

Figure 4: Transformation with reduced extremes

All in all, transformation with reduced extremes leads to high correspondence with categorical scales (Figure 5) and is superior to linear transformation.

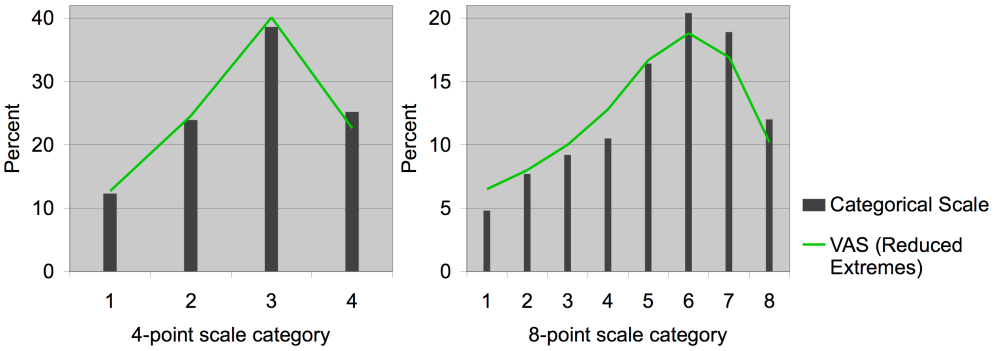

Figure 5: Frequencies of categorical scales and VAS after transformation with reduced extremes

In turn this means that the exteme categories of categorical scales represent a smaller intervall than the other categories; they are less frequently used by the respondents. The centers of categories and categorical scales have different distances, they are not perceived as equidistant.

Experiment 2 – Modifying Categorical Scales

To compensate for the findings of the first experiment that extreme categories represent smaller intervalls of intensity, we modified the categorical scales in the second experiment. The space between the extreme categories and the adjoining ones was scaled down (Figure 6). We expected these changes to influence participants' preference for extreme categories.

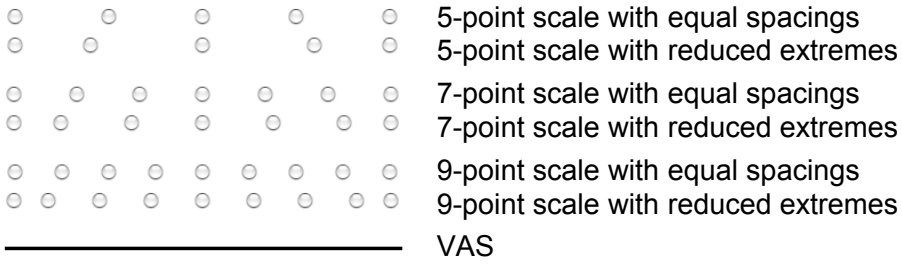

Figure 6: Categorical scales with different spacings and VAS used in experiment two

The participants (n=185) were randomized to one of these seven different scales to rate 58 items on 29 pages. The frequencies of scales with the same number of categories were compared to a VAS that was categorized linearly (Figure 7) to the respective number of categories.

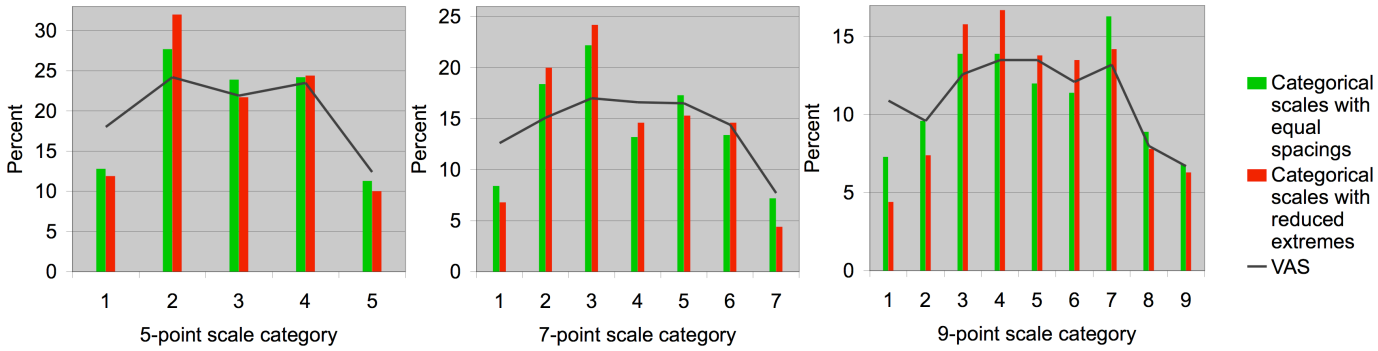

Figure 7: Frequencies of categorical scales with equal spacings, categorical scales with reduced extremes and linearly transformed VAS

The manipulation of the extreme spacings had a clear influence on the observed frequencies. This effect is independent of the number of categories and even the slight divergence between the two variations of the 9-point scale made a difference.
When using categorical scales with reduced extremes, the frequencies of extreme categories decreased.

Conclusion

A general conclusion that can be drawn is that it makes a difference for the distribution of frequencies whether items are rated by categorical scales or by VAS.
In both experiments, the categorization of VAS according to the model of reduced extremes is superior to linear transformation. This means: if one wants to compare frequencies of data obtained from VAS with categorical data, linear transformation is inappropriate and categorization with reduced extremes should be applied.
As modification of spacings between categories has shown to produce robust effects that result in changes in the distribution of frequencies, upcoming research will focus on experiments with spacings between categories. In a further experimental design, the influence of categorical scales with extended extremes will be examined (Figure 8).

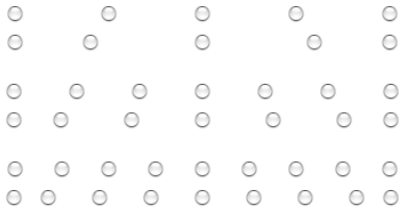

Figure 8: Categorical scales with extended extremes and VAS

References:

Hayes, M. H. S. & D. G. Paterson (1921): Experimental development of the graphic rating method. *Psychological Bulletin, 18*, 98-99

Flynn, D., P. van Schaik & A. van Wersh (2004): A Comparison of Multi-Item Likert and Visual Analogue Scales for the Assessment of Transactionally Defined Coping Function. *European Journal of Psychological Assessment, 20*, 49-58

Nyrén, O. et al., (1987): Self-rating of pain in non-ulcer dyspepsia. *Journal of Clinical Gastroenterology, 9*, 408-414

Schmidt L. C. & I. H. Frieze (1997): A mediational model of power, affiliation and achievement motives and power involvement. *Journal of Business and Psychology, 4*, 425-446