Experimental Determination of Quantifiers for Usability Questionnaire Design

Richard Bates and Robert Bierton

HCI Research Group, Department of Computing Science
De Montfort University, The Gateway, Leicester, UK

Email: [rbates, rbierton]@dmu.ac.uk

ABSTRACT

This paper presents experimentally derived perceived scale positions and distributions of a set of questionnaire quantifiers for use in subjective usability and workload questionnaires. It highlights potential problem quantifiers and presents optimal quantifiers for 5, 7, 9 and 11 interval questionnaire scales.

KEYWORDS: questionnaire design, workload assessment, usability testing.

INTRODUCTION

A survey of usability and workload questionnaires used in HCI work has shown that there appears to be a large variation in questionnaire design with no one common quantifier set used for these questionnaires. Quantifier sets for frequency and amount have been produced previously, Bass et al. (1974), Schriesheim et al. (1978). These sets are, however, not entirely suitable for the type of questions found in usability or workload questionnaires. For example, the set produced by Bass had typical recommended interval quantifiers for seven-interval scales that comprised: 'All', 'An extraordinary amount of', 'A great amount of', 'Quite a bit of', 'A moderate amount of', 'Somewhat', 'None'. It is clear that the Bass quantifiers are not suitable for usability and workload questionnaires. For example, questions such as 'How difficult was the device to use?' cannot reasonably have the choices 'None' or 'All'.

GENERATION OF QUANTIFIER SETS

The method of magnitude estimation, Schriesheim et al. (1978) was chosen for the experiment with a set of 21 candidate quantifiers produced by panel discussion. The keywords 'Happy' and 'Sad' were used with the quantifiers to produce simple and understandable questionnaires. In the experiment 50 postgraduate students and staff were asked to indicate where on a bi-polar line (figure 1) they felt each of the 21 quantifiers should be placed. Responses were scored on a –100, 0, +100 scale.

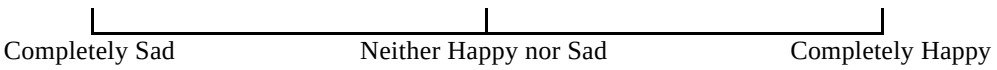

Figure 1. Quantifier test response line

The results were ranked by the 50^{th} percentile of the responses, with the 25^{th} and 75^{th} percentiles giving the inter-quartile ranges for each quantifier (figure 2).

CHOICE OF QUANTIFIERS

The reader may select questionnaire quantifiers for any number of scale intervals by choosing the closest quantifiers to the desired intervals for that scale from figure 2. When a choice of quantifiers is available, the quantifier with the minimum overlap to the adjacent chosen quantifiers and with the smallest inter-quartile range should be selected. Using this method, quantifiers are presented for 11, 9, 7 and 5 interval bi-polar questionnaire scales (figure 3).

Quantifiers	*Percentiles*			
	50th	*25th*	*75th*	*IQD*
Completely happy	*100*	-	-	-
Extremely happy	93.3	90.0	96.7	6.7
Very happy	80.0	73.3	86.7	13.3
Greatly happy	80.0	70.0	86.7	16.7
Really happy	76.7	66.7	86.7	20.0
Very much happy	76.7	63.3	86.7	23.3
Considerably happy	66.7	60.0	76.7	16.7
Happy	48.3	40.0	66.7	26.7
Pretty much happy	46.7	30.0	66.7	36.7
Not at all sad	41.7	0.0	70.0	70.0
Fairly happy	38.3	30.0	50.0	20.0
Quite happy	36.7	26.7	50.0	23.3
Moderately happy	36.7	20.0	43.3	23.3
Somewhat happy	26.7	13.3	36.7	23.3
Just happy	25.0	6.7	43.3	36.7
A little happy	13.3	10.0	20.0	10.0
Slightly happy	13.3	10.0	16.7	6.7
A bit happy	13.3	6.7	20.0	13.3
Not very sad	10.0	-10.0	26.7	36.7
Very slightly happy	6.7	3.3	10.0	6.7
Scarcely happy	3.3	-3.3	10.0	13.3
Neither Happy nor Sad	*0.0*	-	-	-
Scarcely sad	-10.0	-16.7	0.0	16.7
Very slightly sad	-10.0	-20.0	-6.7	13.3
A bit sad	-20.0	-26.7	-10.0	16.7
Slightly sad	-20.0	-26.7	-16.7	10.0
A little sad	-20.0	-30.0	-13.3	16.7
Not very happy	-20.0	-43.3	3.3	46.7
Somewhat sad	-30.0	-40.0	-20.0	20.0
Just sad	-40.0	-56.7	-13.3	43.3
Not at all happy	-40.0	-76.7	-3.3	73.3
Moderately sad	-50.0	-56.7	-36.7	20.0
Quite sad	-51.7	-63.3	-36.7	26.7
Fairly sad	-51.7	-63.3	-36.7	26.7
Sad	-56.7	-66.7	-50.0	16.7
Pretty much sad	-61.7	-73.3	-46.7	26.7
Considerably sad	-73.3	-83.3	-63.3	20.0
Very much sad	-86.7	-90.0	-76.7	13.3
Greatly sad	-86.7	-90.0	-80.0	10.0
Really sad	-86.7	-93.3	-76.7	16.7
Very sad	-90.0	-93.3	-83.3	10.0
Extremely sad	-96.7	-100.0	-93.3	6.7
Completely sad	*-100*	-	-	-

Figure 2. Quantifier test results

11 interval	*9 interval*
Extremely	Extremely
Really	Considerably
Happy	Happy
Moderately	Somewhat
Slightly	Neither/Nor
Neither/Nor	Somewhat
Slightly	Sad
Moderately	Considerably
Sad	Extremely
Really	
Extremely	

7 interval	*5 interval*
Extremely	Extremely
Considerably	Happy
Somewhat	Neither/Nor
Neither/Nor	Sad
Somewhat	Extremely
Considerably	
Extremely	

Figure 3. Optimal quantifiers

DISCUSSION

The importance of using clearly understandable and un-ambiguous quantifiers is shown by the high variability of the 'Not at all' and 'Not very' quantifiers when compared to the variability found in other quantifiers. There is a high degree of symmetry about the 'Neither/Nor' centre point with only small inconsistencies in the ordering of quantifiers. This shows that the subjective meanings of the quantifiers are stable when applied to both positive and negative keywords and hence that valid, symmetrical bi-polar scales can be derived from the data.

CONCLUSIONS

This paper provides a data set of quantifiers suitable for subjective usability and workload evaluation questionnaires. Further work will be carried out to investigate the variation in ratings of certain quantifiers and to generate a further set of quantifiers for a uni-polar, 'Not Happy' to 'Completely Happy' scale.

REFERENCES

Bass, B. M., Cascio, W. F. (1974) Magnitude estimations of expressions of frequency and amount. *Journal of Applied Psychology* **59**(3), 313-320.

Schriesheim, C.A., Schriesheim, J.F. (1978). The invariance of anchor points obtained by magnitude estimation and pair-comparison. *Educational and Psychological Measurement* **38**, 977-983.