Development of a Health Rating scale - Emoticon Health Rating Scale (EHRS) – Selection of items

Integrated Study Report on Study 1:1 and 1:2

Thorsten Klint

2016-07-18

This report reflects the two first studies in a basic development program divided in three phases.

Innehåll

Integrated Study Report on Study 1:1 and 1:2

Development of a Health Rating scale - Emoticon Health Rating Scale (EHRS) – Selection of items

Introduction

Stress related illness constitute an overwhelming majority of sick leave in the Swedish working population. Needless to say the cost for society, employer and employee is enormous.
The overall aim of the current project is to develop a fast, reliable, and web-based health rating scale with the purpose to monitor the respondent's health, giving them the possibility for self-reflection and thereby preventing or minimizing stress related illness. An analysis module will be couple to the instrument for further data analysis of the respondent's health. This module will also include a system indicating of an individual's improvements or deterioration in health (e.g. person's entering into a risk zone for burnout).

Investigators and Study Administrative Structure

Study Sites

The study was conducted by Nextconsulting in cooperation with PFM Research

Sponsor

The study was financed and administered by MYHAPPINESS.SE/Ginsam AB & Sverker Månsson

Project Director

Thorsten Klint, Ph.D.; Associated Professor

Project Statistician

Göran Granath, Ph.D.

Date of study initiation

Study 1:1 February 2016
Study 1:2 April 2016

Date of study completion

Study 1:1 May 2106
Study 1:2 July 2016

Study Objectives

Primary Objectives

To select variables/questions from the literature reflecting different dimensions in the spectrum of health versus illness. These variables are submitted to a web-based panel in order to find out which variables are correlated and thus give similar information. That in contrast to variables which are not correlated and thus could possibly give other types of information. The purpose of the first phase is mainly to reduce the number of variables assessed useful for the final rating scale, by empirical testing in a panel of test persons, and subsequent analysis by multivariate statistics.

Study Design Overview and Rationale

Study Design

Our aim is to reflect health from a salutogenic perspective and thereby using the WHO´s original definition of health "Health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity."

The intention was to select variables reflecting different aspects of health, like mood, sleep, stress and physical activity, from validated rating scales in the literature. A major problem during the selection was that a large majority of rating scales are concerned with a disease perspective. Despite that we were able to identify a number of domains reflecting health and the kind of variables likely to reflect health from a salutogenic perspective. A major source in the identification of rating scales was *Measuring health* by McDowell (2006), and about 50 other references were also consulted.

In the process, to develop a web-based health rating scale for self- monitoring, we first selected about 40 variables (questions) from the literature (only 35 were used in the first study, because 5 variables were missing for technical reasons) and 47 in the second study; original variables plus some extra variables to cover domains we though were of special interest. The next step was to test the selected variables in a web-based panel to find out how the variables were correlated and clustered.

Study Rationale

A telephone recruited web panel from Nextconsulting was used, where each participant had to rate his/her health status at one occasion. The participants were randomly selected from a pool of 70 000 persons living in Sweden, with a representative distribution of the Swedish population concerning age, sex, education, geography etc.

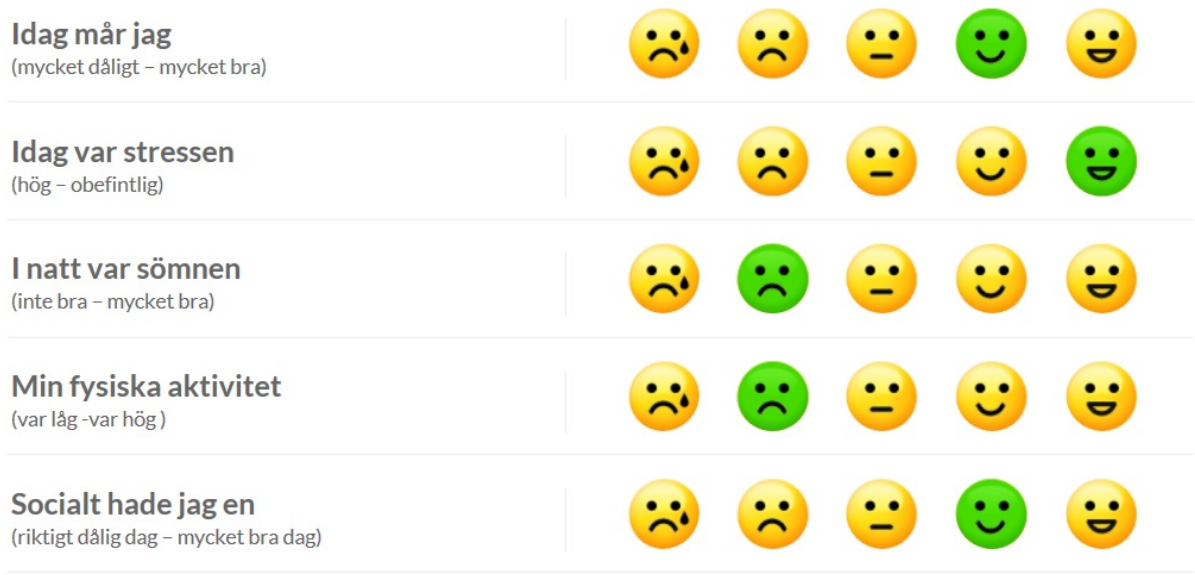
http://www.nextconsulting.se

Variables for Evaluation

The purpose of this study was to determine which variables are correlated and give similar information and which variables are not correlated and could thus possibly give other types of information. The purpose of the first step is mainly to reduce the number of variables useful for the final rating scale, by means of multivariate statistics. The variables used are described in Table 1 and 2. Baseline information about perceived general health, age, gender, education, working situation, and physical exercise were collected.

Variable Layout

The variable layout was designed as in fig 1.



Study Population

Sample Size

No formal sample size calculation was performed, but the number of subjects intended for enrollment was set to 3 times the number of variables (questions in the questionnaire).

Statistical Methodology

Baseline Comparability

Descriptive statistics of age, gender, education etc. was done in order to get a general view of the populations in the two studies.

Data Analysis

The general way of analysis was first to study the correlations between the variables, *ie* the questions and the background data (age, gender etc.). After that PCA and factor analysis was done. The factor analysis was based on PCA and completed with 'Varimax' rotations. That was done in order to find meaningful underlying factors. Non-hierarchical cluster analysis, in variable space, was used in order to group the variables in the number of clusters obtained from the factor analysis. In each cluster, the variable closest to the cluster centre, was chosen. These 'type variables' were the subject to different analyses like correlations and PCA.

Handling of Missing Data

Software

All statistical programming with plots was done using version 13 of the STATISTICA software produced by Dell.

Disposition of Subjects

In study 1:1 121 participants were included. 16 participants never finished the questionnaire and 3 of them scored "I don't know" in more than three question and was subsequently disregarded from the analysis.

In study 1:2 125 participants were included. 20 participants never finished the questionnaire and 3 of them scored "I don't know" or nothing in more than three question and was subsequently disregarded from the analysis.

Results

Primary Objectives Study 1:1

In the first place all 35 variables were subject to a factor analysis. Three factor loadings were isolated which explained about 21 % of the total variance (Table 3). Thereafter the variables were Varimax-rotated and three new factors isolated. Out of the 35 variables factor-loadings > 0.65 were used as a cut-off, which generated 14 variables (Table 4). These variables were then used in a PCA (Table 5). At least three variable groupings were identified. A tentative labelling of the groups could be: 1. Stress variables, 2. Mood variables and "feeling good" variables. The fourth grouping consists of variables reflecting "meaning of life" (Fig 1). The same grouping was confirmed by a hierarchical cluster analysis (Wards method) (Fig 2).

Primary Objectives Study 1:2

In the first place all 47 variables were subject to a factor analysis. Three factor loadings were isolated which explained about 29 % of the total variance. Thereafter the variables were Varimax-rotated and three new factors isolated. All variables were then used in a PCA (Fig 3). A number of tentative variable groups could be suggested, but the grouping was not very distinct since a large number of variables very highly correlated. Therefor a none-hierarchical cluster analysis was used as an alternative strategy. 9 clusters were used in the final analysis, but with variable Q32.2 (Variable 6 in Table 1; *erotic feelings*) excluded since it was it was significantly different between genders. Nine variables were selected based on the importance in the cluster analysis (Fig 4). Indicator variables are marked with yellow in Table 5. The same grouping was also illustrated by hierarchical cluster analysis (Wards method) (Fig 5).

Comparisons between study populations

Using the common 35 variables the distribution of background variables was compared between the two studies. There were no significant differences concerning gender, age, working situation and sleep, however, there were significant differences concerning levels of education and physical exercise. The number of participants completing gymnasium was higher in the second study but the number of participants with > 2 years university education was highest in the first study. Comparing the variables (questions) between the two studies only two variables were significantly different (analysis of variance). Variable Q37.2 (15) "*Today I feel stressed*" and Q37.5 (18) "*Today I have had time for myself*" differed significantly between study populations.

Discussion

The overall aim of the first phase was mainly to reduce the number of variables useful for development of a fast and reliable rating scale reflecting a subject's general health versus illness. This was done by letting persons from a web-based panel scoring the variables at one moment in time to find out which variables potentially were correlated or divided in clusters possibly giving different kind of information.

The first study showed that a large number of variables, as expected, were correlated but also that clear clusters emerged based on the PCA and cluster analysis. Interesting and important was that the correlations and groupings seemed to make common sense.

In the second study the number of variables was increased from 35 to 47 to cover more aspect of the health spectra and compensate for 5 variables missing in the first run. In the first PCA most variables were very well correlated except *erotic feelings* (Q32.2) and the two *stress-related variables* (Q27.3 and 37.2). Q27.1 *physical activity* could possibly also be considered to separate from the rest in p2. Both the none-hierarchical and hierarchical cluster analysis and gave a basis for nine useful variables to be used in further testing. To be noticed the question about erotic feelings was not included in this analysis.

To want extent the differences in background variables *education* and *physical exercise* between the studies reflect on differences in the variables *"Today I feel stressed" and* "*Today I have had time for myself*" is difficult to evaluate.

Conclusions

The overall aim to reduce the number of variables from the original 35 and 47 was substantiated by the two studies. The analysis gives a firm ground for further studies with approximately 9 variables or more. The data shows reasonable similarity between studies and the clustering of variables describes domains of health versus illness which make intuitive sense.

References

Antonovsky A. Health, Stress and Coping. San Francisco: Jossey-Bass; 1979

Antonovsky A. Unraveling the Mystery of Health. How People Manage Stress and Stay Well. San Francisco: Jossey-Bass; 1987.

Lundgren-Nilsson et al. Internal construct validity of the Shirom-Melamed Burnout Questionnaire (SMBQ). BMC Public Health 2012, 12:1

McDowell I. Measuring Health: A Guide to Rating Scales and Questionnaires, OXFORD UNIVERSITY PRESS, Third Edition 2006 (Book)

Nutbeam, D. Health promotion glossary. Health Promotion, 1, 113–127, 1986.

Smith B. J. , K. C. Tang and D. Nutbeam. WHO Health Promotion Glossary: new terms. Health Promotion International, Vol. 21 No. 4, 2006.

Tables and Figures

Table 1 - Variables used in study 1:1

1. Idag var jag fysiskt aktiv (stämmer inte alls – stämmer väldigt bra)
2. I natt kändes min sömn bra (stämmer inte alls – stämmer väldigt bra)
3. Idag känns arbetspressen väldigt stor (stämmer inte alls – stämmer väldigt bra)
4. Idag känner jag mig uppskattad (stämmer inte alls – stämmer väldigt bra)
5. Idag känns det lätt att fatta beslut (stämmer inte alls – stämmer väldigt bra)
6. Idag känner jag erotiska känslor (stämmer inte alls – stämmer väldigt bra)
7. Idag känner jag att jag har kontroll (stämmer inte alls – stämmer väldigt bra)
8. Idag njuter jag av livet (stämmer inte alls – stämmer väldigt bra)
9. Idag är jag mån om hur andra mår (stämmer inte alls – stämmer väldigt bra)

10. Idag känner jag mig snabb i tanken (stämmer inte alls – stämmer väldigt bra)
11. Idag känns problem lätta att hantera (stämmer inte alls – stämmer väldigt bra)
12. Idag har jag haft tid att slappna av (stämmer inte alls – stämmer väldigt bra)
13. Idag känns mina kontakter med andra givande (stämmer inte alls – stämmer väldigt bra)

14. Idag känns balansen mellan arbete och fritid bra (stämmer inte alls – stämmer väldigt bra)
15. Idag känner jag mig stressad (stämmer inte alls – stämmer väldigt bra)
16. Idag känner jag arbetsglädje (stämmer inte alls – stämmer väldigt bra)
17. Idag känner jag mig respekterad (stämmer inte alls – stämmer väldigt bra)
18. Idag har jag haft tid för mig själv (stämmer inte alls – stämmer väldigt bra)

19. Idag känner jag mig mentalt (kraftlös - energisk)
20. Idag känner jag mig (ledsen -glad)
21. Idag känner jag att mitt självförtroende är (dåligt - gott)
22. Idag känner jag mig (olycklig - lycklig)

23. Idag känner jag mig fysiskt (svag - stark)
24. Idag upplever jag livet (meningslöst – meningsfullt)
25. Idag känner jag mig (missnöjd - nöjd)
26. Idag har jag en känsla av (ensamhet – samhörighet)

27. Idag känner jag mig (nere – uppåt)
28. Idag känner jag mig (nervös – avslappad)
29. Idag känner jag mig (utsliten – fräsch)
30. Idag känner jag mig (deprimerad – glad)

31. Idag känner jag att mitt arbete är (meningslöst - meningsfullt)
32. När jag vakande kände jag mig (trött - utsövd)
33. Idag känner jag mig (rädd – orädd)
34. Idag känner jag mig (improduktiv - produktiv)
35. Idag känner jag mig (bekymrad – bekymmerslös)

36. Idag känns mina kontakter med andra (meningslösa – givande)
37. Idag känner jag mig (orolig – lugn)
38. Idag känner jag mig (splittrad – fokuserad)
39. Idag känner jag mig (utanför – omtyckt)
40. Idag var jag (stressad – lugn)

Table 2 - Variables used in study 1:2

1	Q27.1	Idag var jag fysiskt aktiv (stämmer inte alls – stämmer väldigt bra)
2	Q27.2	I natt kändes min sömn bra (stämmer inte alls – stämmer väldigt bra)
3	Q27.3	Idag känns arbetspressen (liten – stor)
4	Q27.4	Idag känner jag mig uppskattad (stämmer inte alls – stämmer väldigt bra)
5	Q32.1	Idag känns min förmåga att fatta beslut (svårt – lätt)
6	Q32.2	Idag känner jag erotiska känslor (stämmer inte alls – stämmer väldigt bra)
7	Q32.3	Idag känner jag att jag har kontroll (stämmer inte alls – stämmer väldigt bra)
8	Q32.4	Idag njuter jag av livet (stämmer inte alls – stämmer väldigt bra)
9	Q87.1	Idag är jag mån om hur andra mår (stämmer inte alls – stämmer väldigt bra)
10	Q87.2	Idag känns mina tankar (långsamma – snabba)
11	Q87.3	Idag känns mina problem (svåra att hantera – lätta att hantera)
12	Q87.4	Idag har jag haft tid att slappna av (stämmer inte alls – stämmer väldigt bra)
13	Q87.5	Idag känns mina kontakter med andra givande (stämmer inte alls – stämmer väldigt bra)
14	Q37.1	Idag känns balansen mellan arbete och fritid bra (stämmer inte alls – stämmer väldigt bra)
15	Q37.2	Idag känner jag mig stressad (stämmer inte alls – stämmer väldigt bra)
16	Q37.3	Idag känner jag arbetsglädje (stämmer inte alls – stämmer väldigt bra)
17	Q37.4	Idag känner jag mig respekterad (stämmer inte alls – stämmer väldigt bra)
18	Q37.5	Idag har jag haft tid för mig själv (stämmer inte alls – stämmer väldigt bra)
19	Q43.1	Idag känner jag mig mentalt (kraftlös - energisk)
20	Q43.2	Idag känner jag mig (ledsen -glad)
21	Q43.3	Idag känner jag att mitt självförtroende är (dåligt - gott)
22	Q43.4	Idag känner jag mig (olycklig - lycklig)
23	Q48.1	Idag känner jag mig fysiskt (svag - stark)
24	Q48.2	Idag upplever jag livet (meningslöst – meningsfullt)
25	Q48.3	Idag känner jag mig (missnöjd - nöjd)
26	Q48.4	Idag har jag en känsla av (ensamhet – samhörighet)
27	Q53.1	Idag känner jag mig (nere – uppåt)
28	Q53.2	Idag känner jag mig (nervös – avslappad)
29	Q53.3	Idag känner jag mig (utsliten – fräsch)
30	Q53.4	Idag känner jag mig (deprimerad – glad)
31	Q58.1	Idag känner jag att mitt arbete är (meningslöst - meningsfullt)
32	Q58.2	När jag vaknade kände jag mig (trött - utsövd)
33	Q58.3	Idag känner jag mig (rädd – orädd)
34	Q58.4	Idag känner jag mig (improduktiv - produktiv)
35	Q58.5	Idag känner jag mig (bekymrad – bekymmerslös)
36	Q64.1	Idag känns mina kontakter med andra (meningslösa – givande)
37	Q64.2	Idag känner jag mig (orolig – lugn)
38	Q64.3	Idag känner jag mig (splittrad – fokuserad)
39	Q64.4	Idag känner jag mig (utanför – omtyckt)
40	Q64.5	Idag känner jag mig (stressad – lugn)
41	Q99.1	Idag känner jag mig (irriterade – behärskad)
42	Q99.2	Idag känns kraven från omgivningen (betungande – obefintliga)
43	Q99.3	Idag känner jag mig (okoncentrerade – koncentrerad)
44	Q99.4	Idag känner jag mig (fysiskt utmattad – full av energi)
45	Q93.1	Idag rörde jag mig regelbundet under dagen (stämmer inte alls - stämmer väldigt bra)

46	Q93.2	Idag kändes mina arbetsuppgifter (väldigt svåra – väldigt lätta)
47	Q93.3	Idag mådde jag (riktigt dåligt – riktigt bra)
48	Q93.4	Idag tog jag regelbundna pauser och slappnade av (stämmer inte alls - stämmer väldigt bra)

Table 3 - Study 1:1

Factor analysis, unrotated, 3 significant factors.

Variable	Factor Loadings (Unrotated) (Klusres2) Extraction: Principal components (Marked loadings are >.500000) Factor 1	Factor 2	Factor 3
Q27.1	-0.47452	0.306841	-0.226525
Q27.2	-0.30760	0.352775	-0.041350
Q27.3	-0.09699	0.103792	0.717767
Q27.4	-0.61914	0.394705	-0.019355
Q32.1	-0.66163	0.497543	0.013014
Q32.2	-0.22598	0.411381	-0.106573
Q32.3	-0.53905	0.427287	0.172105
Q32.4	-0.63065	0.448223	0.135069
Q37.1	-0.57853	0.228394	0.084715
Q37.2	-0.35306	0.112608	0.676306
Q37.3	-0.58030	0.455772	0.022988
Q37.4	-0.56564	0.143192	0.291954
Q37.5	-0.08625	0.169157	0.398202
Q43.1	-0.76902	0.103783	-0.132069
Q43.2	-0.73058	-0.000165	0.046456
Q43.3	-0.73927	-0.030412	0.017662
Q43.4	-0.79521	-0.342072	-0.040188
Q48.1	-0.75651	0.031843	-0.237818
Q48.2	-0.82037	-0.260049	-0.097736
Q48.3	-0.83492	-0.220230	-0.081725
Q48.4	-0.76022	-0.228507	-0.123089
Q53.1	-0.85825	-0.106211	-0.074141
Q53.2	-0.76594	-0.331218	0.170325
Q53.3	-0.82775	0.116376	-0.113114
Q53.4	-0.87539	-0.142776	-0.049717
Q58.1	-0.70262	0.252297	-0.139975
Q58.2	-0.45998	0.430275	-0.003057
Q58.3	-0.61121	-0.451389	0.018865
Q58.4	-0.70382	0.078525	-0.368328
Q58.5	-0.78612	-0.276928	0.201047
Q64.1	-0.77055	0.242895	-0.157475
Q64.2	-0.83144	-0.336429	0.089437
Q64.3	-0.85237	-0.195713	0.005818
Q64.4	-0.77039	-0.315546	-0.004153
Q64.5	-0.74452	-0.369249	0.285908
Expl.Var	15.96380	2.941897	1.806751
Prp.Totl	0.45611	0.084054	0.051621

Table 4 - Study 1:1

Factor analysis; Varimax-rotated; 35 variables.

Variable	Factor Loadings (Varimax raw) (Klusres2) Extraction: Principal components (Marked loadings are >.500000) Factor 1	Factor 2	Factor 3
Q27.1	0.23175	0.543908	-0.145226
Q27.2	0.06545	0.464395	0.028802
Q27.3	0.01929	0.031137	0.730772
Q27.4	0.30342	0.663745	0.082933
Q32.1	0.28265	0.767070	0.131119
Q32.2	-0.03444	0.478745	-0.035744
Q32.3	0.21712	0.619015	0.269168
Q32.4	0.28270	0.691258	0.243113
Q37.1	0.35941	0.488442	0.162132
Q37.2	0.22933	0.182619	0.713289
Q37.3	0.23723	0.687125	0.128793
Q37.4	0.39359	0.379914	0.355593
Q37.5	-0.02302	0.127259	0.421774
Q43.1	0.58864	0.520146	-0.050612
Q43.2	0.61185	0.386594	0.109948
Q43.3	0.63586	0.370500	0.078576
Q43.4	0.85328	0.150892	-0.011467
Q48.1	0.61821	0.469550	-0.164982
Q48.2	0.83001	0.241044	-0.056229
Q48.3	0.82035	0.279510	-0.034300
Q48.4	0.76259	0.238572	-0.082748
Q53.1	0.77760	0.385476	-0.010948
Q53.2	0.82128	0.112689	0.195429
Q53.3	0.63083	0.559399	-0.025197
Q53.4	0.81174	0.360749	0.010276
Q58.1	0.45199	0.608674	-0.046262
Q58.2	0.15054	0.605042	0.089407
Q58.3	0.75835	-0.047703	0.017554
Q58.4	0.54955	0.499352	-0.292989
Q58.5	0.80832	0.163984	0.234150
Q64.1	0.51417	0.640095	-0.058750
Q64.2	0.87959	0.155738	0.120559
Q64.3	0.82094	0.296167	0.056758
Q64.4	0.81774	0.154133	0.025200
Q64.5	0.82323	0.052357	0.303229
Expl.Var	12.08277	6.697522	1.932160
Prp.Totl	0.34522	0.191358	0.055205

Figure 1 - Study 1:1

PCA plot of p1 vs p2

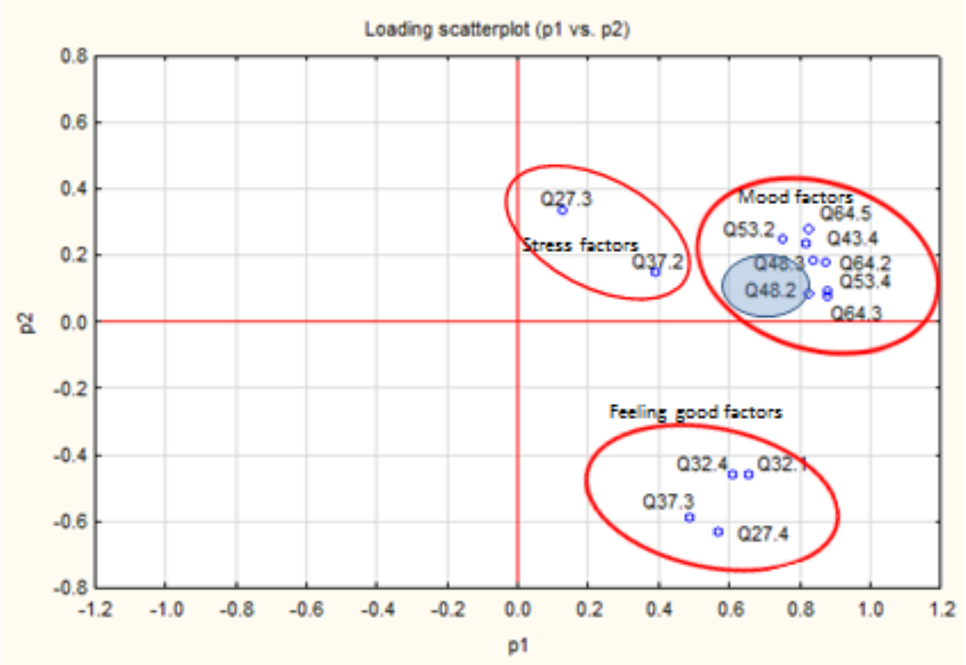


Figure 2 - Study 1:1

14 variables compared with a hierarchical cluster analysis

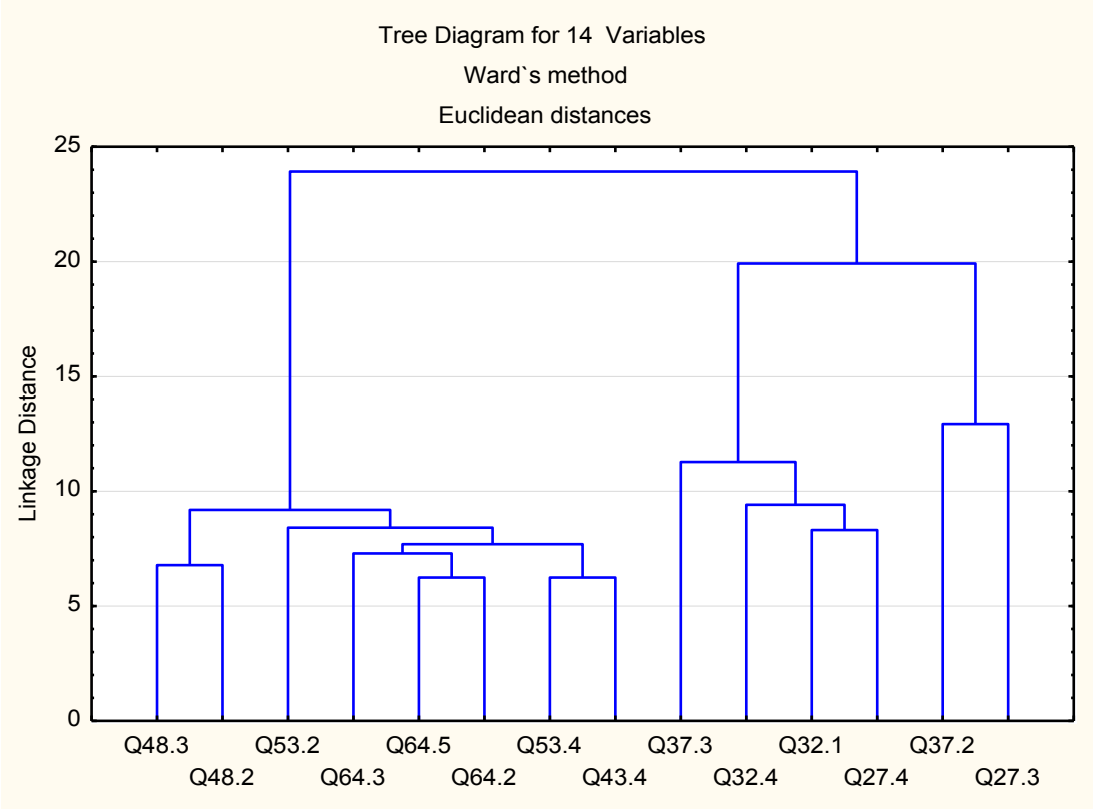


Figure 3 - Study 1:2

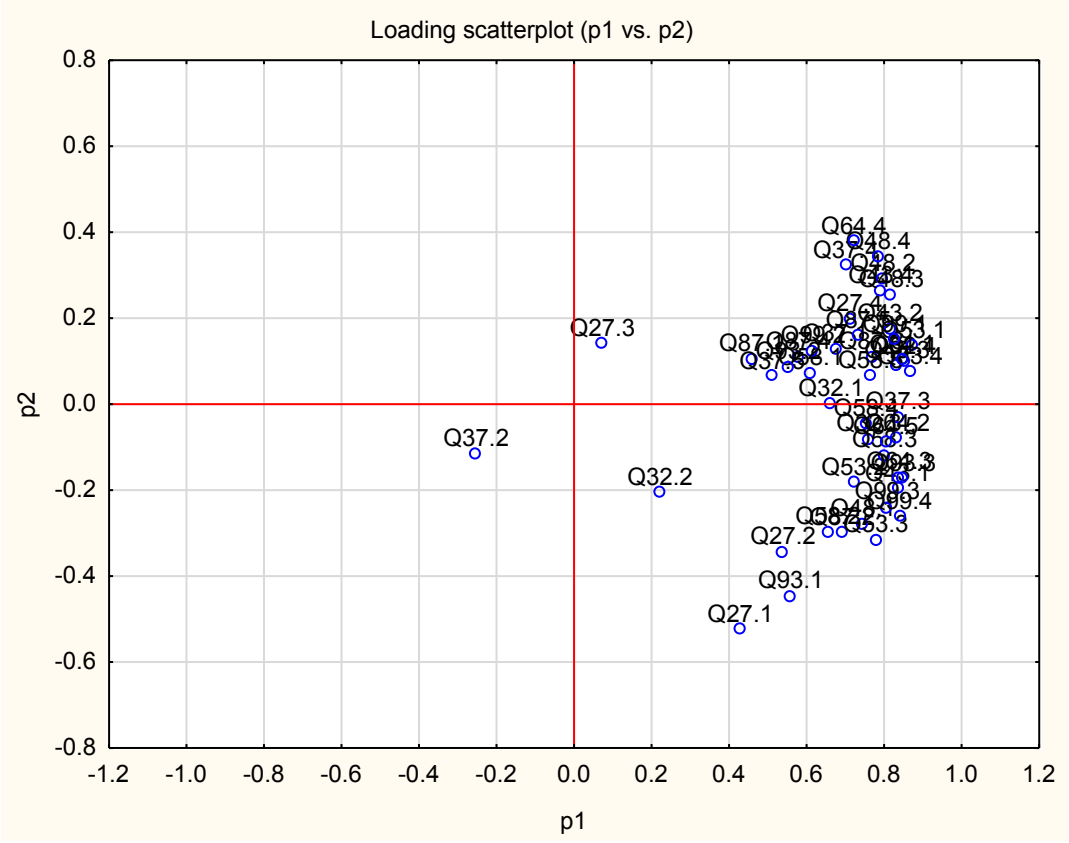


Table 5 - Study 1:2

None-hierarchical	analysis with 9 clusters		
	VARIABLE	CLUSTER	DISTANCE
Q93.3	49	1	0,45
Q64.3	40	1	0,47
Q99.4	46	1	0,49
Q99.3	45	1	0,51
Q43.1	21	1	0,55
Q64.2	39	1	0,56
Q53.3	31	1	0,59
Q64.5	42	1	0,59
Q37.3	18	1	0,60
Q53.2	30	1	0,62
Q87.2	12	1	0,64
Q48.1	25	1	0,65
Q58.2	34	1	0,95
Q53.1	29	2	0,37
Q43.4	24	2	0,46
Q48.3	27	2	0,50
Q43.2	22	2	0,51
Q53.4	32	2	0,54
Q58.5	37	2	0,54
Q43.3	23	2	0,55
Q87.3	13	2	0,61
Q58.4	36	2	0,62
Q64.4	41	3	0,45
Q64.1	38	3	0,48
Q48.4	28	3	0,53
Q48.2	26	3	0,56
Q58.1	33	3	0,70
Q37.5	20	4	0,61
Q87.4	14	4	0,64
Q37.1	16	4	0,71
Q93.4	50	4	0,77
Q37.2	17	5	0,00
Q32.3	9	6	0,53
Q32.4	10	6	0,55
Q87.5	15	6	0,55
Q37.4	19	6	0,55
Q27.4	6	6	0,56
Q27.3	5	7	0,00
Q99.1	43	8	0,53
Q58.3	35	8	0,59
Q32.1	7	8	0,61
Q87.1	11	8	0,70
Q99.2	44	8	0,71
Q93.2	48	8	0,74
Q27.2	4	8	0,79
Q27.1	3	9	0,54
Q93.1	47	9	0,54

Figure 4 - Study 1:2

Variable importance

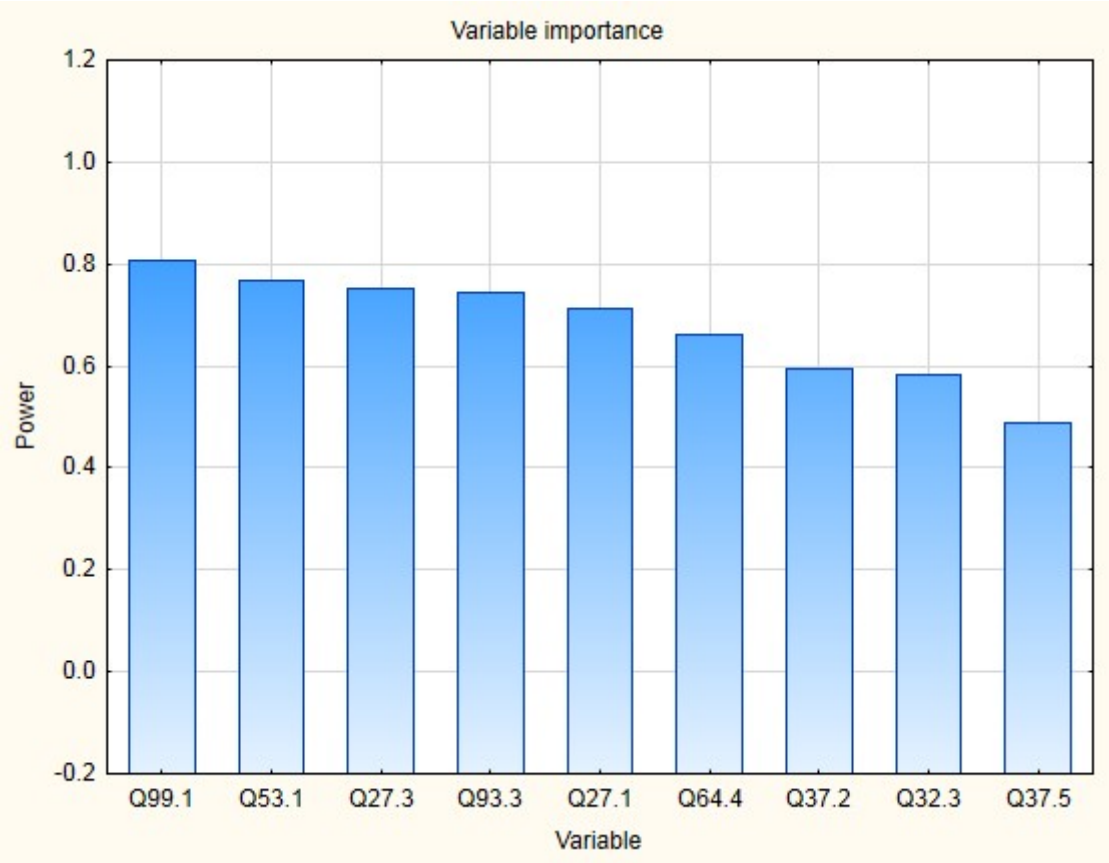
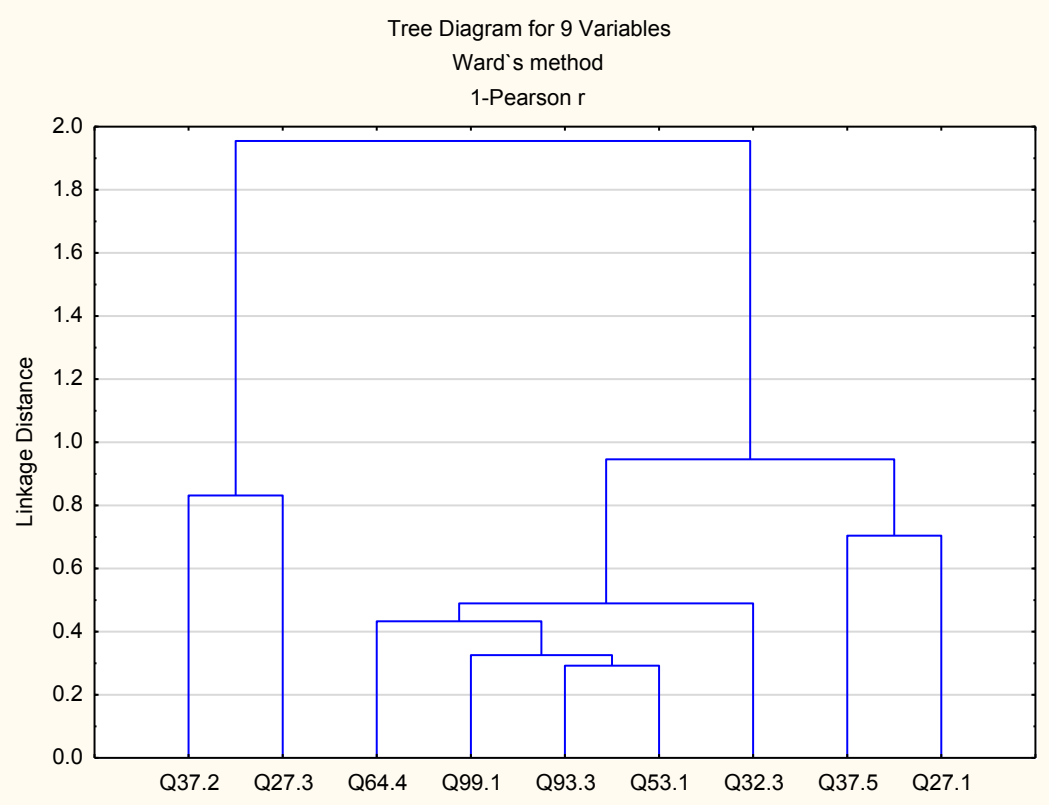


Figure 5 – Stud 1:2


Tree Diagram for 9 Variables
Ward`s method
1-Pearson r
Linkage Distance
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Q37.2
Q27.3
Q64.4
Q99.1
Q93.3
Q53.1
Q32.3
Q37.5
Q27.1

Signaturers

Sponsor

…….
Sverker Månsson, MBA

Project Director

.....
Thorsten Klint, Ph.D.; Associated Professor

Project Statistician

.....
Göran Granath, Ph.D.