Conceptual Modeling of Self-Rated Intelligence-Profile

Petri Nokelainen and Henry Tirri
Helsinki University of Technology Finland,

Kirsi Tirri and Erkki Komulainen
Finland University of Helsinki, Finland

The purpose of this study is to explore the psychometric quality of a self-evaluation instrument based on Gardner's theory of multiple intelligences (1991; 1993; 1995). Self-evaluated intelligence is closely related to a person's self-concept and can reflect both general and academic components of it. Our data includes Finnish students from five different universities (N=256). The questionnaire consists of 70 items that were operationalized from Gardner's theory. The students used a 7-point Likert scale to assess their strengths in each of seven intelligence areas. We performed both exploratory factor analysis and Bayesian network modeling to produce the most plausible structure for self-rated intelligence. The confirmatory factor analysis was conducted in order to study structural relations of self-rated intelligence. We present the 7-component model and discuss the advantages and disadvantages of the final selection of components.

Gardner's Theory of Multiple Intelligences

Gardner's theory of multiple intelligences builds on a concept of an "intelligence", which he defines as "the ability to solve problems, or to create products, that are valued within one or more cultural settings" (Gardner, 1993, x). Considering this definition, Gardner lists seven intelligences that meet his criteria for intelligence. These intelligences are: (1) linguistic, (2) logical-mathematical, (3) musical, (4) spatial, (5) bodily-kinesthetic, (6) interpersonal, and (7) intrapersonal (Gardner, 1993, xi).

Sternberg (1985) identifies Gardner's theory of multiple intelligences as a systems approach similar to his own triarchic theory. Although he likes Gardner's assessments at a theoretical level, he believes them to be a psychometric nightmare. The biggest challenge for advocates of Gardner's approach is to demonstrate the psychometric soundness of their instrument. Sternberg is calling for hard data that would show that the theory works operationally in a way that will satisfy scientists as well as teachers. Twenty years ago quite the same situation was present in the field of self-concept research (e.g., Burns, 1979; Shavelson, Hubner, & Stanton, 1976). However, the situation has changed since methodologically oriented self-concept researchers begin to apply advanced statistical methods to develop robust measurement instruments (e.g., Marsh & O'Neill, 1984; Marsh, Balla & Hau, 1996).

Sternberg's own theory promises the broader measurement implied by the triarchic theory (Sternberg, 1985). His theory provides process scores for componential processing, coping with novelty, automatization, and practical-contextual intelligence, and content scores for the vermatization, and quantitative, and figural content domains (Sternberg, 1991, 266). In the educational setting Sternberg's theory can be used as a framework in planning a program that would meet the needs of different learners (Tirri, 1997). Gardner has shown a special interest in the school's possibilities and limitations to encourage different talents in students (Gardner, 1991). Gardner's theory has been applied in educational settings and in schools (see, e.g., Armstrong, 1993). Gardner warns against using his theory as the only educational approach. There is no single way to adapt his theory, but he has given some guidelines for the possible uses of his theory in schools (Gardner, 1995, 206-209).

Concerning the nature of virtual studying and the experiences of Virtual Open University students in Finland (Tirri & Nevgi, 2001; Nokelainen, Silander, Tirri, H., Nevgi, Tirri, K., 2001), we argue that it is necessary to create self-assessment tools for teachers and students. In the IQ FORM project (Niemi & Ruohotie, 2002), Gardner's theory is used as a guiding theory to build tools for students' self-evaluation (Tirri & Komulainen, 2002).

Self-evaluated intelligence is closely related to a person's self-concept (SC). According to leading researchers, self-concept has a two-factor structure: general self-concept and academic self-concept (Shavelson et al., 1976). Byrne & Gavin (1996) argue that SC is a multidimensional construct, which in their study comprised the four facets of general, academic, English, and mathematics self-concepts. Self-evaluated intelligence can reflect both general and academic components of a person's self-concept. Furthermore, self-evaluated intelligence is closely related to a person's self-esteem and self-confidence. The concept of self-efficacy needs to be acknowledged in the context of self-evaluation. According to Bandura (1978), self-efficacy is specific to a particular activity or situation, in contrast to global beliefs like self-concept. In our study, we concentrate on the self-evaluated intelligence within the Gardnerian framework. We assume that students reflect both general and academic self-concepts in their self-assessments of their strengths and weaknesses.

Methodological Approach and Results

Selection of the Reliable Items

In this paper, our aim is to explore the psychometric quality of a self-evaluation instrument based on Gardner's theory. Our data includes students from five different Finnish universities (N=256). The students represent different disciplines, for example, teacher education, forestry and computer science. Our questionnaire consisted of 70 items that were operationalized from Gardner's theory. Ten items per intelligence represented each particular area. The students used a 7-point Likert scale to assess their strengths in each of seven intelligence areas. The items used to measure each of Gardner's intelligence, and the items used in the 7 components factor solutions reported in the forthcoming text, are listed in Appendix 1 with the whole instrument (see Appendix).

The validity of a self-evaluation instrument is affected by the same defects as any rating system. In general for rating systems, in addition to halo effects, which are difficult to avoid, the following three types of error are often associated with rating scales: the error of severity ("a general tendency to rate all individuals too low on all characteristics"), the error of leniency, ("an opposite tendency to rate too high"), and the error of central tendency, (a "general tendency to avoid all extreme judgments and rate high down the middle of a rating scale") (Kerlinger, 1973, 548-549). The general response tendency in this study shows that the students' have used all the seven options in their answers. However, if all the items used were stacked into one single column, the distribution of responses into the seven alternatives in the scale can be described as unimodal, platycurtic and negatively skewed. Means (mean level of all items) between subjects (N=256) vary heavily (min=2.77, max=5.86). A two-way mixed-effect ANOVA shows that the between people variation is about 11% of the all-item mean variation. This focuses on the fact that response set and/or general self-esteem is strongly present in these measurements. The between measures (P=70) (within people) variation is also quite notable (min=2.25, max=5.50), this share being almost 15%. (Table 1.)

The items with the lowest means (for example 19, 51, 6) refer to specific actions, like writing little songs or instrumental pieces, keeping a diary, or forming mental pictures of objects by touching them. All these activities are so specific that it is natural that the students have given low ratings on them. The items with the lowest standard deviations (for example 44, 28, 35) are such that they do not discern the student population very well. This can be explained by the nature of the items. Most of them measure general attitudes or talents that are needed in academic life. The items refer to the tendency to look for consistency, models and logical series in things, realistic idea of person's strengths and weaknesses and the ability to teach others something you know yourself. The highest rated items include items related to self-reflection and social skills (for example 3, 16, 59). This is very natural because university students need the ability to understand their own feelings and motives in order to plan their academic studies. Furthermore, even in academic studies co-operation and teamwork are necessary elements of successful learning. (Table 1.)

The second phase of item analysis is methodologically multi-staged, starting from the correlations and closing with the results of MIMIC-type modeling. The question is whether the inter-item covariances (correlations) could be reasonably well-conceptualized using Gardner's seven intelligences (or their derivates). The analysis begins by having a look at the correlation matrix. There are 2415 correlation coefficients when the diagonal and double-presentations are omitted. Their mean is 0.11 and range from -0.74 to +0.81.

From the previous two-way mixed effect model ANOVA, it is clear that the average measure intra-class correlation (identical to all-item alpha) is 0.90. The measurement area could be treated as only one dimension: the level of self-evaluated intelligence. However, one can find rather independent components in such a seemingly homogenous set of items. Some deviating correlations cause problems in methods that will be used. The high negative r=-.74 between items 1 and 70 is especially disturbing and a clear outsider in the distribution. There is also a positive tail, correlations above say +.6. This fact points to many, but rather specific, components in the matrix. The correlative properties of items with other items also differ quite a lot. The column (and row) means of correlations and their dispersion properties (to the other 69 items) show clearly that there are items, which cannot be part of any substantive concept/factor. The same phenomenon can be seen in a condensed way in the initial communality values of the items (squared multiple correlation, SMC), and also in the way they load to the first principal component. (Table 2.)

The two first criteria will omit the items 12, 11, 6, 60 and 57. The third criteria omit the item 1, which refers to the school experiences in math, physics and chemistry. This item is not well formulated and it is prone to errors. A person might be very good in mathematics, physics and chemistry, but still not rank them as his/her favorite subjects. Multitalented people might enjoy more, for example, arts and physical education and rate item 1 low for these reasons. Liking and being good at something measure different things. Furthermore, the order of being the first rating item might have an influence on the strict rating behavior practiced by students. (Table 2.)

The analysis, however, is started with the full set of items. In the process towards the final though in this phase rather tentative, conclusion of the nature of inter-item structure, useless items are dropped mainly in phases E2, C2 and C3 (Table 3). At the end of this process, 28 items are in use in the 7-component model. We approach the problem by using exploratory (EFA) and confirmatory (CFA) factor analyses. After coming to the tentative view of a plausible number of dimensions, we then add some background variables to the analysis. In the EFA-approach, this is usually done through correlative means, mainly regression analysis. In the CFA-approach, the final step is done using SEM-modeling (the so-called MIMIC-modeling, see Jöreskog & Goldberger, 1975; Loehlin, 1998) and manifest variables with estimated latent factor scores. We use Bayesian approach (Bernardo & Smith, 2000) to dependence modeling in this study to find the model of the probabilistic dependences between all the variables. The best model is the one with a high probability. When the model of dependences is extracted, it can be used to create different scenarios about causalities that might cause them. Besides revealing the structure of the domain of the data, we interactively study the dependency model by probing it. The approach is summarized in Table 3.

Initial Modeling of the Factor Structure

In the following, as we apply EFA, CFA, MIMIC and Bayesian modeling to the data. In each step of the statistical analysis we refer to the cell of the Table 3 when selectively reporting some interesting findings. The first step in the analysis (C1 in Table 3) was the joining of exploratory factor analysis, confirmatory factor analysis and multiple regression models into a MIMIC model. The procedure is clearly depicted in Bijleveld & van der Kamp (1998).

Chi-Square Test of Model Fit had a value of 5791.16, df=2324, p<.001 and Chi-Square/df=2.49. CFI was 0.580 and TLI 0.564. RMSEA (Root Mean Square Error Of Approximation) was estimated to be 0.076, CI90 0.074<->0.079. SRMR (Standardized Root Mean Square Residual) had a value of 0.116. Due to modest fit indices further modeling was not possible without having guidelines from the exploratory side and thus the following step was then to proceed to E1. (Table 3.)

The scree plot of the eigenvalues (of the principal components extracted from the 70*70 correlation matrix) demonstrated, as expected, that there was a strong first component. The number of eigenvalues 1 or greater was 18, after 12 components the curve dropped a little and settled down to a linear decline. However, the scree line did not suggest any clear cut point. The EFA solutions (step E1 in Table 3) are ML (Maximum Likelihood) solutions with both orthogonal and oblique rotations. Some items do not share any common component. The 7-factor solution did not reflect much of the aimed seven components.

In step B1 we investigated probabilistic dependences between all of the variables (for variable description see Appendix) with a Bayesian search algorithm (Myllymäki, Silander, Tirri & Uronen, 2001) in order to find a model with the highest probability. During the extensive search, 4.027.597 models were evaluated. Graphical visualization of Bayesian network contains two components: (1) observed variables visualized as ellipses and (2) dependences visualized as lines between nodes. Solid lines indicate direct causal relations and dashed lines indicate dependency where it is not sure if there is a direct causal influence or latent cause. Variable is considered as independent of all other variables if there is no line attached to it.

First column of Table 4 presents Bayesian causal model of the variables operationalizing Gardner's theory. The seven intelligence's construct quite clear clusters to the network. Each cluster is labeled according to the following list: (1) linguistic, (2) logical-mathematical, (3) musical, (4) spatial, (5) bodily-kinesthetic, (6) interpersonal, and (7) intrapersonal intelligence. It is noticeable that at this point there is no clear indication (i.e., solid line) of pure causal relationship between any variable in the model.

Second and third columns of Table 4 present importance ranking of the variables in the model. The comparison was conducted by slightly changing the final model for each dimension by removing causal relationships between the variables. If the removal made the model less probable, the causal relationship was considered as a strong dependency, and if the removal made the model more probable, the causal relationship was considered as a weak dependency. As we at this initial modeling phase need to gather evidence to reject or accept variable to the final model, only the weakest dependences are listed from strongest to weakest. The figure in "Probability ratio" column indicates probability ratio if the dependency is removed from the final model. Table 4 shows that there exist three classes among the weakest variables (see Appendix): (1) Low (47, 28, 44, 56, 31, 68, 33, 21, 43, 35, 18, 64, 08, 63, 48), Middle (46, 16, 09, 10, 60), and High (39, 52, 07, 04, 13, 59, 36, 54, 70).

The conclusion of the Bayesian analysis is quite the same as found in steps C1 and E1: The scales has to be purified from items, which do not fit to the structure.

Examination of Multidimensionality

The purification of the model started with steps E2 and C2 (Table 3). Only two areas were rather homogeneous as such, namely musical and interpersonal. We present here the musical scale with all the 10 items, as an example of a homogeneous scale and intrapersonal as an example of a scale that could be split into two components (see Figure 1).

In the case of musical scale, only one item, 8 ("It is easy for me to repeat correctly a musical theme from TV, or some other tune."), is not among the homogeneous set. Musical talent can be seen as a rather unidimensional component and thus it has quite a high alpha and T&V reliability (see Table 6). The intrapersonal scale can be better conceived as a two-dimensional concept. The main point is that there are four items that represent factor 1 rather well. There are five items, which belong to factor 2 rather clearly, and item 16 has a substantial loading only on the first principal component. This finding suggests that the number of dimensions should be more that seven. The optional 12-component (53 items) version is discussed exhaustively by Tirri & Komulainen (2002).

Bayesian dependency modeling was applied to each of the seven dimensions in order to compare the strengths of loadings between variables (step B2 in Table 3). Table 5 lists the importance ranking of three dimensions of Gardner's model. Strongest dependences are listed first, weaker ones are listed with the figure indicating probability ratio if this dependency is removed from the final model. The results support the preceding conclusion of the homogeneity of musical dimension due to fact that it contains no isolated variables. (Table 5.)

Interpretation of the Selected Factor Structure

The results of the factor structure analysis with the reliability estimates are presented in Table 6 (steps E3 and C3 in Table 3). The principle of convergent and discriminant validity was guiding the working. The estimates of reliability applied in this study were the Cronbach's alpha (Cronbach, 1951) and the Tarkkonen's reliability for measurement scales (Tarkkonen, 1987). The Cronbach alpha's basic assumptions allow one dimensional reliability examination, but the Tarkkonen's reliability also operates in the context of multidimensional models (Vehkalahti, 2000). Reliabilities in the reduced 7*4=28 items version form a rather sufficient set for a screening device. Tarkkonen's unbiased estimates give the same picture. (Table 6.)

The first column of Table 7 presents Bayesian causal model applied to all selected variables operationalizing Gardner's theory (step B3 in Table 3). Twenty-eight selected variables form substantially clearer clusters when compared to the initial network model of seventy variables. The clusters are labeled as seen in previous Table 6. The initial examination of this visualization suggests that variable 65 ("When I read, I form illustrative pictures or designs in my mind.") should be removed from the final model. The network model indicates that musical (3) and bodily-kinesthetic (5) dimensions form two separate and isolated homogenous clusters. Linguistic (1), logical-mathematical (2), spatial (4), interpersonal (6), and intrapersonal (7) intelligences are closely related to each other through causal relations. Closer examination of probability ratios of dependences reveals that also variables 56 ("Metaphors and vivid verbal expressions help me learn efficiently."), 32 ("I make contact easily with other people."), and 48 ("It is easy for me to conceptualize complex and multidimensional patterns.") should be omitted from the model.

An Initial Modeling of Gardner's Theory

Goal rotation, using one core item per scale in a more influential position in rotation, does not add much to the picture obtained with the full-free, non-constrained EFA. When factor scores are calculated from the ML solution (7 factors, promax rotation), they show, however, a rather good fit to the factor scores. 7-component model was factor analyzed in the CFA mode using congeneric thinking: an index had a path from one latent variable only (step E4a in Table 3). In addition, all the time no error covariances were allowed (the error terms were kept uncorrelated), thus, the model was for several sets of congeneric scores (Jöreskog & Sörbom, 1979, 52-54). The general self-esteem or general self-concept is difficult to model, although there were strong indications of such a dimension in the initial data screening. New empirical material is also needed for cross-validation. The question of general factor in self-evaluation requires a closer examination in the future. The results of CFA indicate that the model fits the data well. The ratio of the chi-square to the degrees of freedom (2.43) and the RMSEA (0.08) are small indicating good model fit. (Table 8.)

The values give no definite answer to the basic question of whether Gardner's model can be confirmed in the self-evaluated intelligence. This inspection indicates, however, that we may proceed with such an instrument and its development. Concerning our psychometric testing, Gardner's theory offers a promising background to the revision of self-concept, especially in its academic part.

Some areas in self-evaluated intelligence are explained by the background information of the students step E4b in Table 3). Table 9 presents the following background variables, gender (1=Male, 2=Female); age (date of birth, from 1950 to 1981); mother tongue (matriculation examination score, from 1 to 6); mathematical skills (matriculation examination score, from 1 to 6); and motivation (self-rated score from 1 to 5), and their zero-order correlations to the dimensions in Gardner's theory. The results indicate that gender is a powerful explanation of verbalness ($r=.49$, $p < .001$) as the females have rated their abilities higher than the males. Furthermore, linguistic ability seems to increase with age ($r=.22$, $p < .001$). According to our findings, the older students rated this component significantly higher than their younger colleagues. Good grades in mother tongue in the matriculation examination also explained students' high ratings in this component ($r=.34$, $p < .001$). The females were shown to rate themselves significantly higher than the males in both interpersonal ($r=.29$, $p < .001$) and the intrapersonal intelligence ($r=.45$, $p < .001$). The males perceived their logical-mathematical skills significantly better than their female colleagues ($r=-.27$, $p < .001$). The older age and high motivation for university studies explained statistically significantly skills in the interpersonal and intrapersonal areas. Students who had received good grades in mathematics in their matriculation examination rated their interpersonal skills ($r=-.22$, $p < .001$) significantly lower than their colleagues who had received lower grades. (Table 9.)

The earlier research (Tirri & Komulainen, 2002) studied the possibility of 12-component model with 53 items. The psychometric exploration indicated that the 12-component model would be more valid and appropriate to measure all the different areas of Gardner's intelligences. However, the optimized 7-component model with 28 items, presented in this study, is shorter and more convenient in practice. Furthermore, the 7-component model revealed the same trends as the more detailed 12-component one. (Tirri & Komulainen, 2002.)

Discussion

In this article, we have explored the psychometric soundness of the self-assessment instrument based on Gardner's theory. We have made an effort to create a self-assessment tool that would help teachers and learners to identify their strengths and weaknesses as life-long learners. The self-evaluated data would give the student an idea of "What kind of person am I and what kind of persons are those with whom I am studying?" We have adopted Gardner's view on multiple intelligences as a theory guiding our instrument building.

We have described the methodological approach in detail by presenting the selection of the reliable items and structures for the instrument. We have given an interpretation for the selected factor structure and discussed the substructure of each of the original seven intelligences. We have described an initial modeling of Gardner's theory with 7 components. The correlations of explanatory variables with both models validate our findings.

We have shown that some areas are easier than others to evaluate by self-assessment. The psychometric exploration shows that intra- and interpersonal intelligences have components for which it is possible, and perhaps even valid, to evaluate by self-assessment. Furthermore, musical, linguistic and logical-mathematical abilities have been easy for students to rate with our instrument. Some areas, like body coordination and spatial perception, are more difficult to assess by self-assessment. All these intelligence areas would require a test to measure the actual performance of the person. However, we claim that for our purposes to help the students to acknowledge their strengths and weaknesses and to grow as human beings a self-evaluated intelligence instrument can be an empowering tool. Furthermore, the information obtained with it can serve the students and teachers in several ways, to plan their future teaching-studying-learning sessions.

We have presented a 7-component model of the Gardnerian scales and shown that this model reveals the substructures of the seven intelligences. The initial modeling we have performed serves as the beginning of a continuing development of an instrument. The modeling done this far seems to show some promise in using Gardner's theory as a starting point. However, we need further evaluation and validation with new data sets and possible MTMM comparison between 7 and 12-component models. In this process, our study functions as a guide in modifying some of the items or to revise the instrument to better serve the purposes of self-evaluated intelligence.

About the Authors

Kirsi Tirri, PhD., is a Professor of Education at the Department of Education, University of Helsinki. Her research interests include gifted education, teacher education, moral education and cross-cultural studies.

Komulainen Erkki, PhD., Docent in Education, is a researcher at the Department of Education, University of Helsinki. He has specialized to questions of quantitative methodology and statistics in behavioral sciences. He works as consultant in several projects in educational psychology and psychiatry.

Petri Nokelainen, EdLic., is a Research Scientist at the Complex Systems Computation Group, Helsinki University of Technology. His research interest lies on the study of applied statistical modeling and modern learning environments.

Henry Tirri, PhD., is a tenured Professor in Computer Science at the Department of Computer Science, University of Helsinki and a Visiting Professor at Stanford University. His current research interests are technology for new learning environments, theory of causal modeling and game-theoretic approaches to learning.

Contact Details

Professor Kirsi Tirri and Docent Erkki Komulainen
Department of Education
P.O.Box 39 (Bulevardi 18)
FIN-00014 University of Helsinki, Finland
Email: firstname.lastname@helsinki.fi
Phone: +358 9 191 28042
Fax: +358 9 191 28073

Professor Henry Tirri and Research Scientist Petri Nokelainen
Complex Systems Computation Group
P.O. Box 9800 (Tammasaarenkatu 3)
FIN-02015 Helsinki University of Technology, Finland
Email: firstname.lastname@hiit.fi
Phone: +358 9 850 10521

References

Armstrong, T. (1993). *7 kinds of smart.* New York: A Blume Book.

Bandura, A. (1978). Reflections on self-efficacy. *Advances in Behavior Research Therapy, 1*, 237-269.

Bernardo, J., & Smith, A. (2000). *Bayesian Theory*. New York: John Wiley & Sons.

Bijleveld, C., & van der Kamp, L. (1998). *Longitudinal data analysis. Designs, models and methods*. SAGE Publications: London.

Burns, R. (1979). *The self-concept: Theory, measurement, development and behaviour*. London: Longman.

Byrne, B., & Gavin, D. (1996). The Shavelson model revisited: Testing for the structure of academic self-concept across pre-, early, and late adolescents. *Journal of Educational Psychology, 88,* 215-228.

Gardner, H. (1991). *The unschooled mind.* London: Fontana Press.

Gardner, H. (1993). *Frames of mind.* New York: Basic Books.

Gardner, H. (1995). Reflections on multiple intelligences. *Phi Delta Kappan* 77, (3) November 1995.

Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrica, 16*, 297-334.

Jöreskog, K., & Sörbom, D. (1979). *Advances in factor analysis and structural equation models*. Cambridge, Massachusetts: Abt Books.

Jöreskog, K., & Goldberger, A. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of American Statistical Association, 70*, 631-639.

Kerlinger, F. (1973). *Foundations of behavioral research.* London: Holt, Rinehart & Winston.

Loehlin, J. (1998). *Latent variable models. An introduction to factor, path, and structural analysis.* Third Edition. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Marsh, H., Balla, J., & Hau, K. (1996). An evaluation of incremental fit indices: A clarification of mathematical and empirical processes. In G.A. Marcoulides & R.A. Schumacker (Eds.), *Advanced structural equation modeling techniques* (pp. 315-353). Hillsdale, NJ: Erlbaum.

Marsh, H., & O'Neill, R. (1984). Self-Description Questionnaire III: The construct validity of multidimensional self-concept ratings by late adolescents. *Journal of Educational Measurement, 21,* 153-174.

Myllymäki, P., Silander, T., Tirri, H., & Uronen, P. (2001). Bayesian Data Mining on the Web with B-Course. In N. Cercone, T. Lin and X. Wu (Eds.), *Proceedings of the 2001 IEEE International Conference on Data Mining* (pp. 626-629). IEEE Computer Society Press.

Niemi, H., & Ruohotie, P. (Eds.) (2002). *Theoretical Understandings for Learning in the Virtual University*. Research Centre for Vocational Education and Training, University of Tampere. Saarijärvi: Saarijärven Offset.

Nokelainen, P., Silander, T., Tirri, H., Nevgi, A., & Tirri, K. (2001). Modeling students' views on the advantages of web-based learning with Bayesian networks. In H. Ruokamo, O. Nykänen, S. Pohjolainen, and P. Hietala (Eds.), *Proceedings of the Tenth International PEG Conference: Intelligent Computer and*

Communications Technology – Learning in On-Line Communities (pp. 101-108).

Shavelson, R., Hubner, J., & Stanton, G. (1996). Self-concept: Validation of construct interpretations. *Review of Educational Research, 46*, 407-441.

Sternberg, R. (1985). *Beyond IQ: a triarchic theory of human intelligence*. New York: Cambridge University Press.

Sternberg, R. (1991). Death, taxes, and bad intelligence tests. Editorial. *Intelligence, 15,* 3, 257-269.

Tarkkonen, L. (1987). *On reliability of composite scales. An essay on the structure of measurement and the properties of the coefficients of reliability – a unified approach.* The Finnish Statistical Society. Statistical Studies 7.

Tirri, K. (1993). *Evaluating teacher effectiveness by self-assessment: a cross-cultural study.* Research report 122. Department of Teacher Education. Helsinki: Yliopistopaino.

Tirri, K. (1997). How Finland meets the needs of gifted and talented pupils. *High Ability Studies, 8,* 2, 213-222.

Tirri, K., & Nevgi, A. (2001). Students' views on learning in virtual university. *Resources in Education,* ED 448121.

Tirri, K., & Komulainen, E. (2002). Modeling a Self-rated Intelligence-profile for Virtual University. In H. Niemi and P. Ruohotie (Eds.), *Theoretical Understandings for Learning in Virtual University* (pp. 139-168). Research Centre for Vocational Education and Training, University of Tampere. Saarijärvi: Saarijärven Offset.

Vehkalahti, K. (2000). *Reliability of measurement scales. Tarkkonen's general method supersedes Cronbach's alpha*. The Finnish Statistical Society. Statistical Studies 17.

Appendix 1

The Items Used in the Instrument

	Original Gardner version	Initial version	Final version
1. Linguistic	04, 09, 12, 31, 34, 40, 56, 60, 69, 70	04, 40, 37, 70	04, 40, 56, 70
2. Logical-mathematical	01, 21, 27, 30, 37, 39, 44, 54, 57, 66	01, 30, 39, 44, 27	01, 30, 39, 54
3. Musical	02, 08, 14, 15, 19, 26, 55, 58, 61, 62	02, 14, 15, 19, 55, 61, 62	14, 15, 55, 62
4. Spatial	05, 10, 13, 18, 20, 38, 48, 53, 63, 65	05, 13, 38, 48, 53, 65	05, 48, 53, 65
5. Bodily-kinesthetic	06, 07, 25, 29, 33, 36, 45, 47, 52, 67	07, 25, 29, 33, 45, 67	29, 33, 45, 67
6. Interpersonal	11, 22, 23, 32, 35, 43, 46, 49, 50, 59	23, 32, 49, 50	22, 23, 32, 59
7. Intrapersonal	03, 16, 17, 24, 28, 41, 42, 51, 64, 68	03, 17, 24, 42, 51	03, 16, 17, 42

The Questionnaire

1. At school I was good at mathematics, physics or chemistry.
2. I am good at singing or playing an instrument.
3. I often think about my own feelings and sentiments and seek reasons for them.
4. Writing is a natural way for me to express myself.
5. At school, geometry and various kinds of assignments involving spatial perception were easier for me than solving equations.
6. I have a talent to form a mental picture of objects by touching them.
7. I am very good at tasks that require good coordination.
8. It is easy for me to repeat correctly a musical theme from TV, or some other tune.
9. I enjoy reading demanding novels or classics.
10. Other people say that I am good with colours.
11. One of my strengths is problem solving together with other people.
12. When walking outside, I am good at finding words on signs and posters and making them rhyme.
13. When I think, I can see clear visual images in my mind.
14. After hearing a tune once or twice I am able to sing or whistle it quite accurately.
15. When listening to music, I am able to discern instruments or recognise melodies.
16. I am able to analyse my own motives and ways of action.
17. I spend time regularly reflecting on the important issues in life.
18. I am able to see objects or events that I would like to document on camera or video.
19. I can write little songs or instrumental pieces.
20. I usually find my way, even in unfamiliar places.
21. It is easy for me to use abstract concepts.
22. Even in strange company, I easily find someone to talk to.
23. I get along easily with different types of people.
24. I have opinions of my own and dare to disagree with others.
25. I have good coordination.
26. I have a good singing voice.
27. I can easily measure, classify, analyse or calculate things.
28. I have a realistic idea of my strengths and weaknesses.
29. I am handy.
30. I can work with and solve complex problems.
31. I am good at entertaining myself and others with wordplay and jokes.
32. I make contact easily with other people.
33. I can easily do something concrete with my hands (e.g. knitting and woodwork)
34. It is easy for me to play with word games, for example crossword puzzles.
35. I am good at teaching others something I know myself.
36. I have the strength to participate in extreme physical experiences (e.g. shooting the rabbits, parachuting and mountain climbing).
37. I easily notice lapses of logic in other people's everyday speech or actions.
38. I am good at jigsaw puzzles, picture puzzles and various kinds of labyrinth puzzles.
39. I am good at games and problem solving which require logical thinking
40. I have recently written something that I am especially proud of, or for which I have received recognition.
41. I am able to handle criticism directed against me.
42. I like to read psychological or philosophical literature to increase my self-knowledge.
43. I am the kind of person that neighbours, colleagues or fellow students turn to for advice and instructions.
44. I tend to look for consistency, models and logical series in things.
45. I am good at showing how to do something in practise.
46. I easily recognise other peoples' motives.
47. It is easy for me to imitate other peoples' gestures, facial expressions and ways of moving.
48. It is easy for me to conceptualise complex and multidimensional patterns.
49. It is easy for me to understand other peoples' feelings and moods.
50. I consider myself a leader (or have been called one by other people).
51. I keep a diary or note down happenings of my inner life.
52. I often "talk with my hands" and/or otherwise use body language when talking to someone.
53. I can easily imagine how a landscape looks from a bird's-eye view.
54. Mental arithmetic is easy for me.

55. I can easily keep the rhythm when drumming a melody.
56. Metaphors and vivid verbal expressions help me learn efficiently.
57. I am good at making decisions or predictions from new scientific discoveries.
58. I play a musical instrument or otherwise take part in musical activities.
59. In negotiations and groupwork, I am able to support the group to find a consensus.
60. I have a talent to use concepts or expressions which are not very typical in other people's everyday talk.
61. I quickly recognise a song or piece of music.
62. I notice immediately if a melody is out of tune.
63. I'm good at drawing and designing various kinds of figures.
64. When necessary, I am able to motivate myself, even for unpleasant tasks.
65. When I read, I form illustrative pictures or designs in my mind.
66. I want to present things as logically as possible and give reasons for them.
67. I was good at handicrafts at school.
68. I can handle the emotions caused by serious setbacks.
69. In conversation, I often refer to things that I have read or heard about.
70. At school studies in native language or social studies were easier for me than mathematics, physics and chemistry.

Figure 1:
The Homogeneous Musical (3) and the Two-dimensional Intrapersonal (7) Scale

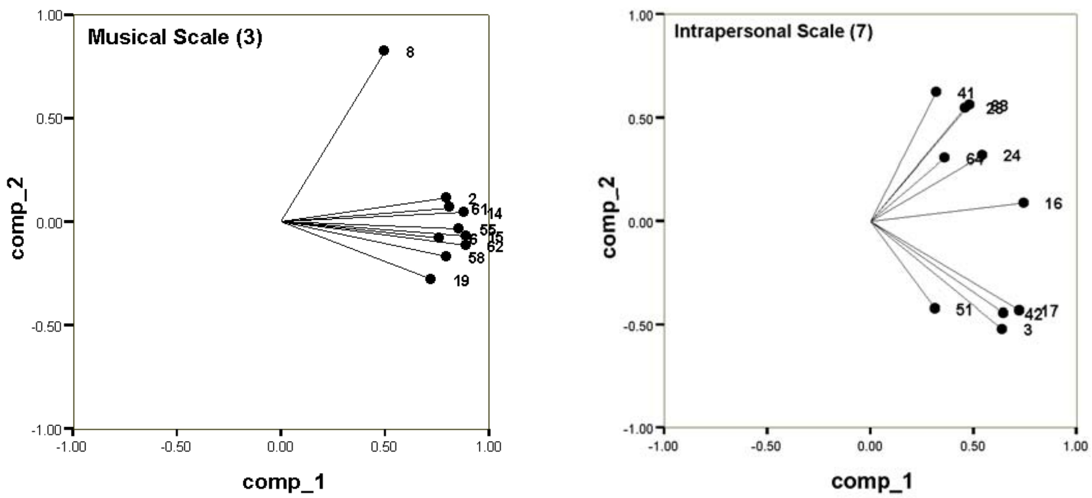


Table 1
Item Level Distributions

		Mean	S.D.	Kurtosis	Skewness
Items stacked (N=17920)		4.46	1.75	-0.84	-0.38
Item means per subject (N=256)	min	2.77	0.92	-1.54	-1.76
	max	5.86	2.29	2.88	0.94
Item means per item (P=70)	min	2.25	1.11	-1.37	-1.13
	max	5.50	2.42	1.69	1.39
Seven items with highest mean (hi -> lo)		3, 16, 59, 13, 37, 35, 29			
Seven items with lowest mean (lo -> hi)		19, 51, 6, 58, 60, 12, 40			
Seven items with highest stdev (hi -> lo)		29, 24, 46, 59, 63, 64, 53			
Seven items with lowest stdev (lo -> hi)		62, 44, 28, 35, 67, 17, 66			

Table 2
The Correlative Properties of Items

	share with other items		
	mean r**2	min	max
	0.036	0.010	0.077
Seven highest (hi -> lo)	14, 15, 32, 62, 55, 61, 2		
Seven lowest (lo -> hi)	12, 11, 60, 6, 69, 64, 57		
	share with other items		
	mean SMC	min	max
	0.592	0.350	0.834
Seven highest (hi -> lo)	62, 15, 14, 32, 33, 29, 25		
Seven lowest (lo -> hi)	12, 6, 20, 11, 57, 60, 64		
	first PC loading	min	max
	0.342	-0.271	0.673
Seven highest (hi -> lo)	32, 49, 14, 22, 15, 26, 2		
Seven lowest (lo -> hi)	1, 54, 39, 27, 12, 38, 67		

Table 3
The Phases in the Examination of the Factor Structure

E1. Exploratory factor analyses using 7 factors with varimax and promax rotation.	B1. Bayesian dependency modeling applied to all variables.	C1. Confirmatory factor analysis according to Gardner's conception and various GFI-estimates.
E2. Each of the seven dimensions analyzed in two dimensions and loading plots.	B2. Bayesian dependency modeling applied to each of the seven dimensions.	C2. Each of the seven dimensions as a congeneric scale.
E3. Reliability estimates (Cronbach's alpha) to the original 7 scales.	B3. Bayesian dependency modeling applied to the selected variables.	C3. Estimates of reliability according to Tarkkonen's procedure.
E4a. Goal/target rotated exploratory factor analysis using the important items by weighting them. E4b. The chosen EFA-model with background variables using factor scores.		C4. The seven-component model in MIMIC with background variables.

Table 4
Initial Bayesian Network Model and the Importance Ranking of the Weakest Loading Variables Measuring Self-Evaluated Intelligence.

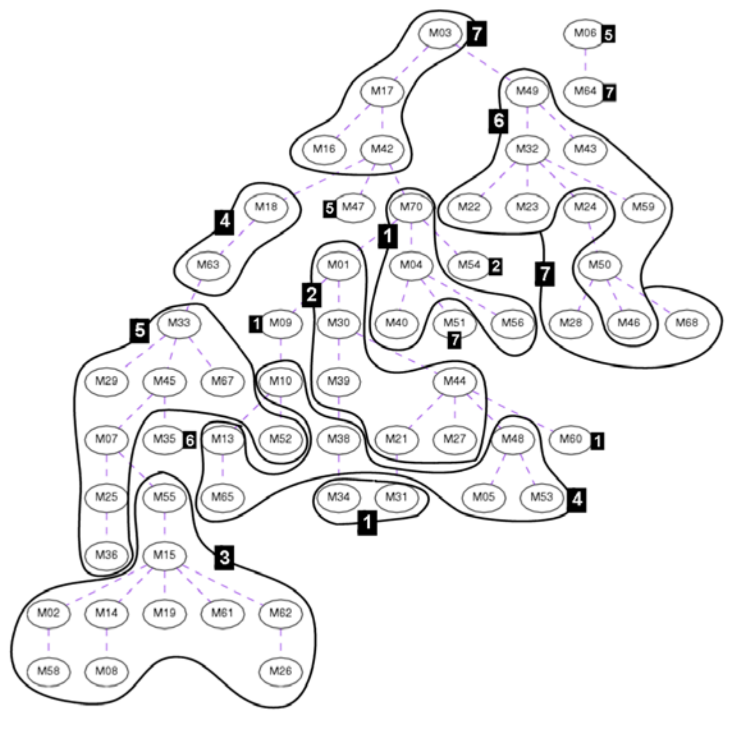
Network Model	Dependency	Probability ratio
	M42->M70	1 : 240265
	M70->M54	1 : 205727
	M25->M36	1 : 162380
	M32->M59	1 : 53296
	M10->M13	1 : 33827
	M70->M04	1 : 31428
	M45->M07	1 : 27590
	M10->M52	1 : 16768
	M30->M39	1 : 12638
	M44->M60	1 : 8373
	M09->M10	1 : 7246
	M01->M09	1 : 6332
	M17->M16	1 : 4761
	M50->M46	1 : 3585
	M44->M48	1 : 892
	M18->M63	1 : 719
	M14->M08	1 : 483
	M06->M64	1 : 238
	M42->M18	1 : 186
	M45->M35	1 : 116
	M49->M43	1 : 106
	M44->M21	1 : 79
	M63->M33	1 : 71
	M50->M68	1 : 34
	M21->M31	1 : 33
	M04->M56	1 : 33
	M30->M44	1 : 32
	M50->M28	1 : 25
	M42->M47	1 : 18

Table 5
Importance Ranking of the Linguistic, Musical, and Intrapersonal Scales in the Bayesian Dependency Model

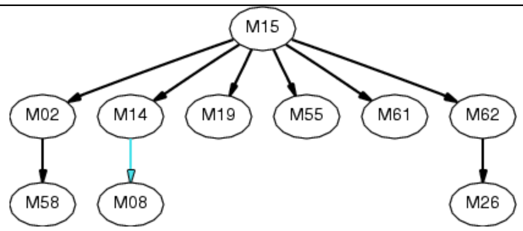
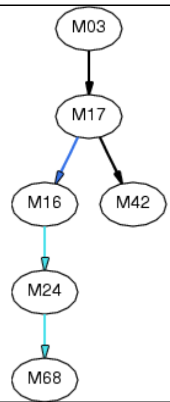
1. Linguistic		3. Musical		7. Intrapersonal	
					
Dependency	Probability ratio	Dependency	Probability ratio	Dependency	Probability ratio
M04 -> M40	1 : Inf.	M15->M61 M15->M14 M15->M55 M15->M62 M15->M02 M02->M58 M15->M19 M62->M26	1 : Inf.	M03->M17 M17->M42	1 : Inf.
M70 -> M04	1 : 35920			M17->M16	1 : 4574
M70 -> M09	1 : 5265			M16->M24	1 : 54
				M24->M68	1 : 24
M12, M31, M34, M60, M69	No importance	M14 -> M08	1 : 324	M28, M41, M51, M64	No importance

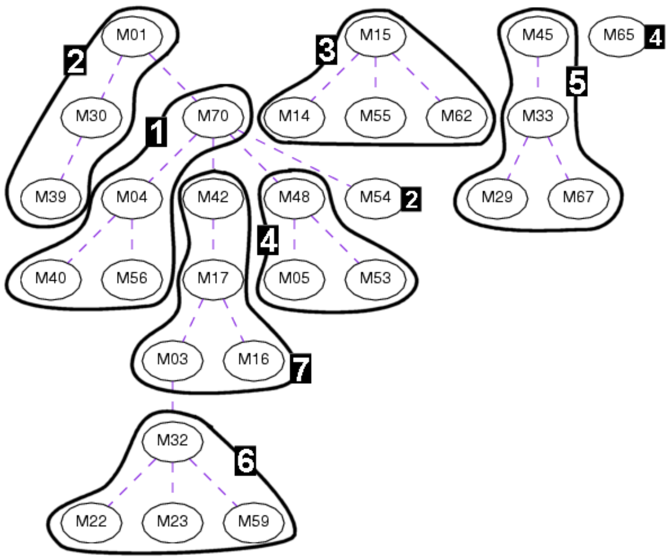
Table 6

The Results of the Factor Structure Analysis with the Reliability Estimates

	Original 7-component (70 items)		Optimized 7-component (28 items)	
Gardner's dimension	Alpha	Tarkkonen	Alpha	Tarkkonen
1. Linguistic	.64	.77	.71	.74
2. Logical- mathematical	.76	.81	.75	.77
3. Musical	.93	.96	.90	.93
4. Spatial	.73	.76	.70	.74
5. Bodily- kinesthetic	.74	.87	.85	.89
6. Interpersonal	.82	.92	.86	.89
7. Intrapersonal	.70	.81	.77	.81

Table 7

Bayesian Network Model and the Importance Ranking of the 28 Selected Variables Measuring Self-Evaluated Intelligence.



Network Model	Dependency	Probability ratio
(see figure)	M33->M29	1 : Inf.
	M01->M70	
	M15->M14	
	M15->M55	
	M15->M62	
	M33->M67	
	M32->M23	
	M32->M22	
	M17->M03	
	M42->M17	
	M01->M30	
	M45->M33	
	M48->M05	1 : 1.000.000
	M04->M40	
	M48->M53	
	M70->M42	1 : 206003
	M70->M54	1 : 177004
	M32->M59	1 : 37933
	M70->M04	1 : 26949
	M30->M39	1 : 10737
	M17->M16	1 : 4042
	M04->M56	1 : 27
	M03->M32	1 : 25
	M70->M48	1 : 11
	M65	No importance

Table 8

The Goodness-of-fit Values for the 7-Component Model

Statistic	7-component optimized model
Chi-Square	799.46
Df	329
p-value	< .001
Chi-square/df	2.43
CFI	0.864
TLI	0.843
RMSEA	0.075
CI90	0.068<->0.081
SRMR	0.088

Table 9
The Background Variables Explaining the Self-Evaluated Intelligence

7-component model (28 items)	Gender	Age	Mother tongue	Math skills	Motivation
1. Linguistic	.49***	.22***	.34***	-.16**	.14*
2. Logical-mathematical	-.27***	-.03	-.02	.19**	.15
3. Musical	.04	.13*	.07	-.07	.13*
4. Spatial	.09	.08	-.06	-.17**	.09
5. Bodily-kinesthetic	.01	.02	-.15*	-.12	.04
6. Interpersonal	.29***	.20***	.07	-.22***	.35***
7. Intrapersonal	.45***	.16**	.05	-.15*	.19**

* $p < .05$ ** $p < .01$ *** $p < .001$