Published paper

Stewart, M.E., Watson, R., Clark, A., Ebmeier, K.P., Deary, I.J. (2010) *A hierarchy of happiness? Mokken scaling analysis of the Oxford Happiness Inventory*, Personality and Individual Differences, 48 (7), pp.845-848
http://dx.doi.org/10.1016/j.paid.2010.02.011

A hierarchy of happiness? Mokken scaling analysis of the Oxford Happiness Inventory.

[a]Mary E. Stewart, [b]Roger Watson, [a]Andrea Clark, [c]Klaus P. Ebmeier, [d]Ian J. Deary

[a] Applied Psychology, School of Life Sciences, Heriot-Watt University, Edinburgh, United Kingdom

[b] School of Nursing, Midwifery & Health, The University of Sheffield, United Kingdom

[c] Department of Psychiatry, University of Oxford, United Kingdom

[d] Centre for Cognitive Ageing and Cognitive Epidemiology, Department of Psychology, University of Edinburgh, United Kingdom

Abstract

The items of the Oxford Happiness Inventory (OHI), a self-report assessment of happiness, are subjected to an analysis for hierarchy among its items. By using Mokken scaling analyses we can assess whether items can reliably be ordered between persons as severity indicators on a latent trait; in this case, a latent trait of Happiness. OHI item-level data from 1024 participants were entered into the Mokken Scaling Procedure (MSP) seeking reliable scales with $H > 0.30$. 12 OHI items formed a reliable and statistically significant hierarchy. However, the MSP values indicate a 'weak' scale. The 'most difficult' (happiest) item on the scale is 'feeling energetic' and the 'least difficult' (least happy) is 'I have fun'. Items in the scale are consistent with what is already known about both happiness and low mood. The reduction in the OHI's items from twenty-nine to twelve in the Mokken scale may have utility making it more accessible to participants as well as identifying items with reliably different levels of 'difficulty'.

Introduction

Research into happiness has increased considerably over the past forty years. A product of this interest in happiness, its causes, and the psychological processes that produce it, has been the development of several self-report scales which aim to measure this construct (Larsen, Diener, & Emmons, 1985; Kashdan, 2004). These subjective measures have allowed progress in happiness research by supplementing early social surveys based on objective variables (Swami, 2008) or 'social' indicators' (Argyle, 2001) of happiness, such as health, income, education, demographics and life events. They have helped to provide a fuller account of happiness (Diener, 1984).

Happiness is generally considered to comprise three main components: the frequency and degree of positive affect or joy; the absence of negative feelings, such as depression or anxiety; and the average level of satisfaction over a period (Andrews & Withey, 1976; Diener, 1984). The first two refer to the affective aspects of the construct, the latter to the cognitive aspects (Diener, Emmons, Larsen & Griffin, 1985). Happiness is, therefore, a multidimensional construct. Some measures of happiness emphasise the affective component (e.g., the Affect Balance Scale, Bradburn, 1969; the Positive and Negative Affect Scale, Watson & Clark, 1988), some emphasise the cognitive component (e.g., the Life Satisfaction Scale, Diener et al., 1985; the Delighted-Terrible Scale, Andrew & Withey, 1976), and others offer a combination as general measures of happiness (e.g., the Oxford Happiness Inventory Argyle, Martin & Crossland, 1989).

Scales measuring happiness range from single questions to multi-item scales. In early research on happiness, researchers tended to rely on a single self-report item to measure each component of happiness (Diener, 2000). Single-item scales, for example, “taking all things together, how would you say things are these days?” (Gurin, Veroff, & Feld, 1960), have been rated using Likert scales, verbal rating scales (e.g., 'very happy', 'pretty happy', 'not too happy'), or by visual analogue scales. Cantril's (1985) 'ladder scale' and Andrew and Withey's (1976) Delighted-Terrible Scale use seven schematic faces whose expressions range along the continuum from very negative to very positive. Although single-item measures can be quite successful for some research investigations (Argyle, 2001), such measures have been criticised in that they cannot be tested for internal consistency reliability (Argyle, 2001; Diener, 1984; Stones, Kozma, Hirdes, Gold & Kolopack, 1996).

Several multi-item measures exist and those that have been widely used include Bradburn's (1969) Affect Balance Scale, the Depression-Happiness Scale (McGreal & Joseph, 1993), Satisfaction With Life Scale (Diener et al., 1985), the Memorial University of Newfoundland Scale of Happiness (Kozma & Stones, 1980, 1983), the Short Happiness and Affect Research Protocol (Stones et al., 1996), the Subjective Happiness Scale (Lyubomirsky & Lepper, 1999), and the Oxford Happiness Inventory (OHI: Argyle et al., 1989).

The present study will assess the hierarchical properties of the items in the OHI, using Mokken scaling. As far as we are aware the OHI has not been subjected to any analysis searching for a hierarchy among its items. A scale with hierarchical properties has items

which are consistently ordered relative to one another, giving an indication of the relative position of each item on the latent trait assessed by the scale; the extent to which each item indicates a 'mild' or 'severe' level of the latent trait. To explain this psychometric concept, consider the following analogy. If weight lifters can successfully lift 130kg, they would not be asked to lift 120kg or 110kg because these weights are easier. Therefore, if, on a scale, a person endorses an item indicating a given level of happiness, they are likely also to have endorsed all items indicating lower levels of the same latent trait. However, this cannot be taken for granted, and the hierarchy of items can be tested empirically.

Hierarchical scales have proven useful in measuring a range of constructs: for instance, Neuroticism (Watson, Deary, & Austin, 2007), psychological distress (Watson, Deary, & Shipley 2008) and feeding behaviour in dementia (Watson, 1996). Such scales enable the presence of a latent trait to be determined, not by finding items' loadings as is done in factor analysis, but by discovering whether and how many specific items form an ordered array along it. The General Health Questionnaire (GHQ) has a number of formats with 60, 30, 28 or 12 items. The GHQ is often used to measure psychological distress and levels of mental ill-health. It was originally developed to assess the extent of psychiatric illness in general practice (Goldberg, 1978). Watson et al. (2008) identified nine items that form a useful and reliable scale by applying Mokken scaling to the GHQ-30. A shorter scale when assessing mental health is of particular clinical utility as it is often difficult for participants to complete long questionnaires.

The Oxford Happiness Inventory (OHI; Argyle et al., 1989) was designed to provide a general measure of personal happiness and followed the design of the Beck Depression

Inventory (BDI; Beck et al., 1961), with a number of the BDI items reversed and, in addition, supplementary items measuring further aspects of subjective well-being (Argyle, 2001). The 29-item OHI follows the same four-choice format as the BDI, with incremental steps defined as follows: unhappy or mildly depressed (e.g., 'I do not feel happy'); a low level of happiness (e.g., 'I feel fairly happy'); a high level of happiness (e.g., 'I am very happy'); and manic (e.g., 'I am incredibly happy') (Francis, 1999). The aggregate score forms the measure of overall happiness, with higher scores indicating greater happiness. Cronbach alpha coefficients for the OHI are high (0.84 to 0.90). However, despite these high alphas for the total scale, factor analysis identifies, in addition to a single factor of happiness, a number of more specific dimensions: satisfaction with life; personal efficacy; sociability/empathy; a positive outlook; physical well-being; cheerfulness; and self-esteem (Argyle, Martin & Lu, 1995; Hills & Argyle, 1998).

The OHI has been widely used in the UK (e.g. Furnham & Brewin, 1990; Joseph & Lewis, 1998) and USA (e.g. Valiant, 1993), and has been used cross-culturally to compare students in Australia, Canada, the UK and USA (Francis, Brown, Lester & Phillipchalk, 1998). There are also several translations of the OHI (see Lewis, Francis & Ziebertz, 2002). The scale shows high internal reliability (e.g. Hills & Argyle, 2001; Argyle & Lu, 1990), with high reliability over time (Argyle et al., 1989; Francis, Ziebertz & Lewis, 2003), and good consensual validity when self-reports are compared with a friend's ratings (Argyle, 2001). It also correlates with measures of all three components of happiness, with large correlations between the OHI and life satisfaction, and with depression: Positive affect (Bradburn positive affect; .32); life satisfaction (Life

Satisfaction Index; .57); negative affect (Bradburn negative affect; -.32: BDI; -.52) (Argyle, Martin & Crossland, 1989).

In the present study, we inquire whether items from the OHI form a hierarchy by using Mokken scaling. We test whether the items can reliably be ordered as severity indicators on a latent trait. If OHI items form a Mokken scale it will help to understand the latent trait of Happiness.

METHODS

Mokken scaling, derived from one of the two branches of item response theory (Hulin, Drasgow, & Parsons, 1983, p. 14) is a stochastic development of Guttman scaling (van Schur, 2003). Both methods seek to establish whether hierarchical scales exist in multivariate datasets. However, whereas Guttman scaling can only process dichotomous responses to items in a questionnaire, Mokken scaling can process multiple response categories. This is enabled through a computer programme: the Mokken Scaling Procedure (MSP; Sijtsma, Debets, & Molenaar, 1990). Furthermore, whereas Guttman scaling is deterministic in nature and very prone to violations by items of the Guttman hierarchy, Mokken scaling, due to its stochastic nature, is more flexible in this respect (Mokken & Lewis, 1982).

Mokken scaling generates the following informative parameters: *scalability* (Loevinger's coefficient or H), *reliability* (Rho; a test-retest statistic analogous to Cronbach's alpha), and *probability* (p, Bonferroni corrected for multiple comparisons) of obtaining a hierarchical scale (Molenaar & Sijtsma, 2000). *Scalability*, in Mokken scaling terms,

refers to the extent to which items in a questionnaire will be ordered hierarchically relative to one another. The Mokken scaling procedure uses the mean scores for endorsement of items to order the items. The Loevinger's coefficient is calculated from a set of items ordered, on the one hand, according to their mean values and, on the other hand, taking into account the number of Guttman violations in the relative ordering of items; the fewer Guttman violations, the greater the scalability. Acceptable values for these criteria have been established as $H>0.3$; $Rho>0.7$, and $p<0.05$.

Procedure

Recruitment of participants

Participants were recruited from Edinburgh's Universities. They were included if English was their first language and if they were over 17 years of age. Participants were given a stamped, addressed envelope to return a questionnaire, and were asked to take a questionnaire only if they were willing to be considered for a later, experimental (acute tryptophan depletion) stage of the study (see Stewart, Deary, & Ebmeier, 2002). Approximately 2000 questionnaires assessing mood and personality were distributed; 1041 questionnaire packs were completed and returned, among which 1036 completed the OHI. The questionnaire pack took approximately 1 hour to complete. Data concerning the following were described previously: the structure of the TPQ (Stewart, Ebmeier, & Deary, 2004); the relationships between personality and mood (Stewart, Ebmeier, Deary, 2005), and whether personality is a predictor of suicidal thoughts (Stewart, Donaghey, Deary, & Ebmeier, 2008).

Data from 1036 participants, reverse-scored for negative items on the OHI, were entered into a SPSS (v15) database. Cases with missing data were deleted listwise and the data saved as a tab-delimited file with the spreadsheet option turned off; 12 participants were excluded by this procedure. These data were imported into the MSP and a search for scales initiated at H=0.05 with this value being increased in 0.05 increments to H=0.55 as recommended by Meijer and Baneke (2001) until reliable scales were obtained above H=0.30. In this case only one reliable scale was obtained at H=0.30 and no scales were obtained above H=0.50. Therefore, the former was used to define a Mokken scale by studying the *'Crit'* values produced by MSP for individual items. *'Crit'* is a statistic generated by MSP combining parameters related to Guttman violations by items, using the H values of all the items remaining in the scale (Molenaar & Sijtsma, 2000). Items with *'Crit'* values lower than 40 are retained in the final scale (Molenaar & Sijtsma, 2000).

RESULTS

Four hundred males, six hundred and nineteen females, and five people who did not indicate their gender, were included in the sample. Ages ranged from 17.1 to 50.4 years (mean = 21.0 yrs, standard deviation = 4.5). Thirteen people did not indicate their date of birth.

The result of the analysis described above is shown in Table 1: a 12-item scale was derived from the original 29 items which has acceptable scalability (H=0.36), is highly statistically significant (p=0.001), and reliable (Rho=0.97). The items run, in terms of ease to difficulty—the extent to which they are adopted by participants—from having fun

through to feeling energetic with a range of other items related to happiness in between. The most 'difficult to achieve' items are related to 'Feeling energetic', and the presence of positive affect 'Feeling happy' and 'Experience of joy', as well as the feeling that 'Life has meaning'. Included within the twelve items are cognitive aspects such as 'I feel that I am in total control of all aspects of my life' and 'I always have a good influence on events'.

Discussion

Twelve items of the 29-item OHI form a reliable hierarchical scale of Happiness. At the level of scalability obtained ($H < 0.4$), the scale would be considered 'weak' as opposed to 'moderate' (Molenaar & Sijtsma, 2000). Nevertheless, the scale obtained is highly reliable and statistically significant. The original items for the OHI were not developed with hierarchical scaling in mind and it is possible that further development of the scale could be undertaken. Obtaining a scale with twelve items from the original twenty-nine demonstrates that, in terms of Mokken scalability, the majority of the original items were redundant. In addition, the reduction in items from twenty-nine to twelve should increase the utility of the scale as shorter scales are easier to complete, being less time consuming and with a lower tendency to confuse respondents through the inclusion of redundant and possibly repetitive items. The ability of Mokken scaling in shortening previously validated scales has recently been demonstrated by Watson et al. (2008) whereby the thirty item version of the GHQ was reduced to nine items forming a reliable and statistically significant Mokken scale which incorporated five items from the previously shortest version of the GHQ the GHQ-12.

The twelve items of the Mokken scale from the OHI are consistent with what is already known about happiness and its reverse. 'Feeling full of energy', 'Feeling Happy' that 'Life has meaning' are particularly important, the opposite of these concepts are often related to negative affect and low quality of life (e.g. WHOQOL Group, 1998). Also within the hierarchy is 'I feel that I am in total control of all aspects of my life', similarly loss of control is often associated with negative affect (e.g. Teasdale, 1983). Warm and joyful feelings and having a positive effect on events and on other people are difficult to obtain, more difficult than happy cognitions, such as memories and the appreciation of beauty. It is perhaps of no surprise that just 'feeling happy' is one of the most difficult items, more difficult than having a purpose and a meaning to life. The only item more difficult than 'feeling happy' is related to feelings of energy and arousal, low energy and fatigue are often associated with depression and low mood. Bradburn (1969) conceptualized happiness as the difference between the scores on positive and negative affect items, from the twelve items included in the hierarchy by our relatively young sample it would seem that the presence of positive affect is more important than the absence of negative items.

The items in the scale relate highly to other factors which have been found to be predictors of happiness. These include social competence (Argyle & Lu, 1990), satisfaction with relationships (Lu & Argyle, 1992), self-esteem (Furnham & Cheng, 2000; Cheng & Furnham, 2003), leisure satisfaction (Lu & Argyle, 1994), and coping styles (Rim, 1993).

There is an obvious advantage and utility of such a condensed hierarchical scale, particularly when assessing individuals who fatigue quickly. In practice, a scale which has hierarchical properties is administered in the same way as any Likert type scale; in theory, the resulting score (in the case of the present scale scores range from 12 to 48 using the 1-4 Likert type rating) should indicate the level of the latent trait present. In addition, this will also provide an indication of the specific aspects of the latent trait that are present. Only items up to that point in the scale should have been endorsed and others above that point should not have been endorsed. Therefore, a score of 4 should indicate that the respondent 'has fun' but nothing else and a score of 48 would indicate that the respondent 'feels energetic' and endorses all the other items below 'feeling energetic' in the hierarchy. Therefore, in theory a score of 36 indicates, in general—bearing in mind the scale's modest H value—that the respondent 'experiences joy' and endorses all the other items below that point but does not, necessarily, feel that 'life has meaning', 'feel happy' and 'feel energetic'. However, no scale is perfect and no set of responses will be perfect and it is possible that some respondents will endorse some items higher up the hierarchy than their scores suggest to varying extents, but these should be rare in a strong and well constructed Mokken scale. In the present study, the scale was not constructed with hierarchical scaling in mind and, clearly, would require further development to establish a stronger hierarchical scale. This study has, nevertheless, demonstrated that the development of such a scale may be possible. The construct validity of the current scale should be tested by measuring against other measures of happiness and positive affect, as well as other predictors of positive mood.

This study was carried out in a group of university students, who were on average in their early twenties. It would, therefore, be of interest to assess whether the same hierarchical scale also holds true across the lifespan. It could be hypothesised that with advancing age the absence of negative items may become more important than the presence of positive items. Being in control may follow an inverted u-shape of financial and practical independence in mid-life.

This study therefore shows that the OHI has hierarchical properties. From previous research we know that the OHI is a reliable and valid measure of happiness. For those in their early twenties at least, we can conclude that feeling energetic, happy and that life has meaning are especially important indicators of a happy life.

References

Andrews, F. M., & Withey, S. B. (1976). *Social indicators of well-being: America's perception of life quality.* New York: Plenum Press.

Argyle, M. (2001). *The Psychology of Happiness (*2nd ed.). Routledge: London.

Argyle, M., & Lu, L. (1990). The happiness of extraverts. *Personality and Individual Differences, 11*(10), 1011-1017.

Argyle, M., Martin, M., & Crossland, J. (1989). Happiness as a function of personality and social encounters. In J. P. Forgas & J. M. Innes (Eds), *Recent advances in social psychology: An international perspective*, (pp. 189-203). North Holland: Elsevier Science Publishers.

Argyle, M., Martin, M., & Lu, L. (1995). Testing for stress and happiness: The role of social and cognitive factors. In Spielberger, C.D. & Sarason, I.G. *Stress and Emotion*, 15, 173-187. Taylor and Francis.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4*, 561-571.

Bradburn, N. M. (1969). *The Structure of Psychological Well-Being.* Chicago: Aldine.

Cantril, H. (1985). *The Pattern of Human Concerns.* New Brunswick, NJ: Rutgers University Press.

Cheng, H., & Furnham, A. (2003). Personality, self-esteem, and demographic predictions of happiness and depression. *Personality and Individual Differences, 34,* 921-942.

Diener, E. (1984). Subjective well being. *Psychological Bulletin, 95,* 542-575.

Diener, E. (2000). Subjective well-being: The science of happiness, and a proposal for a national index. *American Psychologist, 55,* 34-43.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment, 49,* 71-75.

Francis, L. J. (1999). Happiness is a thing called stable extraversion: A further examination of the relationship between the Oxford Happiness Inventory and Eysenck's dimensional model of personality and gender. *Personality and Individual Differences, 26,* 5-11.

Francis, L. J., Brown, L. B., Lester, D., & Philipchalk, R., (1998). Happiness as stable extraversion: A cross-cultural examination of the reliability and validity of the Oxford happiness inventory among students in the UK, USA, Australia, and Canada. *Personality and Individual Differences, 24,* 167-171.

Francis, L. J., Ziebertz, H., & Lewis, C. A. (2003). The relationship between religion and happiness among German students. *Pastoral Psychology, 51,* 273-281.

Furnham, A., & Brewin, C. R. (1990). Personality and happiness. *Personality and Individual Differences, 11,* 1093-1096.

Furnham, A., & Cheng, H. (2000). Perceived parental behaviour, self-esteem and happiness. *Social Psychiatry and Psychiatry Epidemiology, 35*, 463-470.

Goldberg, D. (1978). *General Health Questionnaire*. Windsor: NFER Publishing Company.

Gurin, G., Veroff, J., & Feld, S. (1960). *Americans view of their mental health.* New York: Basic.

Hills, P., & Argyle, M. (1998). Musical and religious experiences and their relationship to happiness. *Personality and Individual Differences, 25*, 91-102.

Hills, P., & Argyle, M. (2001). Emotional stability as a major dimension of happiness. *Personality and Individual Differences, 31*, 1357-1364.

Hulin, C. L., Drasgow, F., & Parsons, C. K. (1983). Introduction to item response theory. In C. L. Hulin, F. Drasgow, & C. K. Parsons (Eds.), *Item response theory* (pp. 13–74). Homewood, Illinois: Dow Jones-Irwin.

Joseph, S. & Lewis, C. A. (1998). The Depression-Happiness Scale: Reliability and validity of a bipolar self-report scale. *Journal of Clinical Psychology, 54*, 537-544.

Kashdan, T. B. (2004). The assessment of subjective well-being (issues raised by the Oxford Happiness Questionnaire). *Personality and Individual Differences, 36*, 1225-1232.

Kozma, A., & Stones, M. J. (1980) The measurement of happiness: Development of the Memorial University of Newfoundland Scale of Happiness (MUNSH). *The Journal of Gerontology, 35,* 906-912.

Kozma, A., & Stones, M. J. (1983) Re-validation of the Memorial University of Newfoundland Scale of Happiness. *Canadian Journal of Aging, 2,* 27-29.

Larsen, R. J., Diener, E., & Emmons, A. (1985). An evaluation of subjective well-being measures. *Social Indicators Research, 17,* 1-17.

Lewis, C. A., Francis, L. J., & Ziebertz, H. (2002) The internal consistency reliability and construct validity of the German translation of the Oxford Happiness Inventory. *North American Journal of Psychology, 4,* 211-220.

Lu, L., & Argyle, M. (1992). Receiving and giving support: Effects on relationships and well-being. *Counselling Psychology Quarterly, 5,* 123-133.

Lu, L., & Argyle, M. (1994). Leisure satisfaction and happiness as a function of leisure activity. The *Kaohsiung Journal of Medical Sciences, 10,* 89-96.

Lyubomirsky, S., & Lepper, H. S. (1999). A measure of subjective happiness: Preliminary reliability and construct validation. *Social Indicators Research, 46,* 137-155.

McGreal, R. & Joseph, S. (1993). The Depression-Happiness Scale. *Psychological Reports, 73,* 1279-1282.

Meijer, R. R., & Baneke, J. J. (2001). Analyzing psychopathology items: A case for nonparametric item response theory modelling. *Psychological Methods*, 9, 354–368.

Mokken, R. J., & Lewis, C. (1982). A nonparametric approach to the analysis of dichotomous item responses. *Applied Psychological Measurement*, 6, 417–430.

Molenaar, I. W., & Sijtsma, K. (2000). MSP5 for Windows. Groningen: iec ProGAMMA.

Rim, Y. (1993). Happiness and coping styles. *Personality and Individual Differences, 14*, 617-618.

Sijtsma, K., Debets, P., & Molenaar, I. W. (1990). Mokken scaling analysis for polychotomous items: Theory, a computer programme and an empirical application. *Quality and Quantity*, 24, 171–188.

Stewart, M.E., Deary, I.J., & Ebmeier, K.P. (2002) Neuroticism as a predictor of Mood Change: The Effects of Tryptophan Depletion. *British Journal of Psychiatry,* 181: 242-247.

Stewart, M.E., Donaghey, C., Deary, I.J. & Ebmeier, K.P. (2008) Suicidal Thoughts in Young People: their frequency and relationships with personality factors. *Personality and Individuals Differences,* 44, 809-820.

Stewart, M.E., Ebmeier, K.P., & Deary, I.J. (2004) The Structure of Cloninger's Tridimensional Personality Questionnaire in a British Sample. *Personality and Individual Differences,* 36, 1403-1418.

Stewart, M.E., Ebmeier, K.P. & Deary, I.J (2005) Personality correlates of happiness and sadness: EPQ-R and TPQ compared. *Personality and Individual Differences*, 38 1085-1096.

Stones, M. J., Kozma, A., Hirdes, J., Gold, D., Arbuckle, T., & Kolopack, P. (1996). Short Happiness and Affect Research Protocol (SHARP). *Social Indicators Research, 37,* 75-91.

Swami, V. (2008). Translation and validation of the Malay Subjective Happiness Scale. *Social Indicators Research, 88,* 347-353.

Teasdale, J.D. (1983) Negative thinking in depression: Cause, effect, or reciprocal relationship. *Advances in Behaviour Research & Therapy*,5, 3-25.

Valiant, G. L. (1993). Life events, happiness and depression: The half empty cup. *Personality and Individual Differences, 15,* 447-453.

van Schur, W. H. (2003). Mokken scale analysis: Between the Guttman scale and parametric item response theory. *Political Analysis*, 11, 139–163.

Watson, D., & Clark, L. A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS Scales. *Journal of Personality and Social Psychology, 54*, 1063-1070.

Watson, R. (1996). Mokken scaling procedure (MSP) applied to feeding difficulty in elderly people with dementia. *International Journal of Nursing Studies,* 33, 385–393.

Watson, R., Deary, I., & Austin, E. (2007). Are personality trait items reliably more or less 'difficult'? Mokken scaling of the NEO-FFI. *Personality and Individual Differences*, 43, 1460–1469.

Watson, R., Deary, I.J., & Shipley, B (2008) A hierarchy of distress: Mokken scaling of the GHQ-30 *Psychological Medicine,* 38, 575–579.

WHOQOL Group (1998) The World Health Organisation Quality of Life assessment (WHOQOL): development and general psychometric properties. *Social Science and Medicine,* 46, 1569-1585.

Table 1

Mokken scale of all Oxford Happiness Inventory (OHI) checked for violations of monotone homogeneity and double monotonicity (n=1024)

OHI item paraphrased	Mean	H
Feeling energetic	2.13	0.34
Feeling happy	2.29	0.42
Life has meaning	2.35	0.41
Experience of joy	2.44	0.49
Good influence on events	2.58	0.35
Cheerful effect on others	2.64	0.38
Warm feelings	2.64	0.32
Happy memories	2.65	0.33
Finding beauty	2.70	0.32
In total control	2.75	0.36
Laugh a lot	2.81	0.38
I have fun	3.02	0.45

H=0.38; Reliability = 0.83; p=0.00041

Mean = 31.00; Standard deviation = 4.67; Skewness = -0.37; Kurtosis = 0.10