Reducing the Length of a

Goldberg Based Personality Questionnaire Using

Item Response Theory & Confirmatory Factor Analysis

A thesis presented in partial fulfilment of the requirements for the degree of

Masters of Science

in

Psychology

at Massey University, Albany,
New Zealand.

Nathan Conrad Phillips

2006

Supervised by
Dr Richard Fletcher

Abstract

Objectives: This study seeks to reconstitute an existing personality questionnaire by identifying the items that capture the best quality information as measured through *Item Response Theory* (IRT). This process will reduce the length of this measure and increase its measurement precision.

Method: A polytomous IRT model (*Graded Response*: Samejima, 1969) will be used to assess the psychometric properties of each item in this questionnaire and produce item level graphs in order to select the best three items for each of the 26 first-order factors. *Confirmatory Factor Analysis* (CFA) will be used to assess the model fit and unidimensionality before and after the IRT selections are made. This will illustrate the improvement gained through both the deletion of redundant items and the selection of high-quality items.

Results: This questionnaire was reduced from 246 items down to 78 items with three high-quality items identified for each of the 26 first-order factors. The model fit considerably improved through this selection process and the reduction of information was minimal in comparison to the amount of items that were deleted.

Conclusions: This study illustrated the power of using IRT for test development. The item selections are not only of benefit for the organisation that supplied the data for this study, but also the original developers as well as any other users of these items as they are freely available via an online source.

Dedication:

I dedicate this thesis to my wife Lara, who has been highly supportive of me throughout this long process.
Lara helped me keep on track and I would not have finished this without her.
This was not an easy topic to get my head around let alone try to explain to friends and family.
Thank you Lara for your love and support at all stages of my life.
I think it is finally time for me to start work!

Table of Contents

List of Tables

List of Figures

List of Appendices

Introduction

Personality Testing

Personality is assessed through determining and measuring individual characteristics or traits that represent important differences between people (Ozer & Reise, 1994). Personality is also viewed as being relatively stable across situations and across time and therefore has many applications if measured in an appropriate manner.

A focal reason for the study of personality stems from the desire to scientifically understand human behaviour. The use of this information is largely of interest to psychologists and other behavioural researchers, but it is also of great interest to organisations. Meta-analyses have illustrated the importance of the relationship between certain personality characteristics and organisational outcomes (Barrick & Mount, 1991; Barrick, Mount, & Judge, 2001). By understanding these relationships organisations are better equipped to seek further information about applicants for positions or promotions and therefore make better decisions.

The Development of Personality Testing

Many questionnaires have been designed and continuously refined to improve the usefulness of their output and the efficiency of their input (Costa & McCrae, 1997). This refinement process began over 60 years ago with the foundation being laid by Raymond B. Cattell (Goldberg, 1990). Cattell was one of the first scientists to apply empirical procedures to the task of constructing a taxonomy of personality items, and achieved this by assessing the correlations amongst the items and by using oblique rotational procedures (Goldberg, 1990).

Cattell (1943) worked to define a short list of categories that encompassed thousands of English personality characteristic adjectives and concluded that the 171 scales he developed could parsimoniously be grouped into a dozen different categories. The academic consensus that followed Cattell's foundation work was that the immense list of items could be grouped under five major headings (Goldberg, 1990).

The Big-Five and Five-Factor Models

The understanding of personality through the measurement of personality traits is widely accepted with the dominant method utilising the five factors alluded to above (Ozer & Reise, 1994). The term applied to this form of grouping is the *Five-Factor Model* (FFM; Guenole & Chernyshenko, 2005) with the most common FFM referred to as the *Big-Five* (Goldberg, 1990). The categories used for tests such as these are traditionally numbered and labelled as follows: (1) *Surgency* or *Extraversion*, (2) *Agreeableness*, (3) *Conscientiousness* or *Dependability*, (4) *Emotional Stability* or *Neuroticism*, and (5) *Culture* or *Intellect* or *Openness* (Goldberg, 1990). These five factors have been shown to account for a large proportion of the variance in self-report personality questionnaires (Guenole & Chernyshenko, 2005) meaning that these fives factors give a good overall impression of an individual's personality. For a full discussion of the history of the *Big-Five* see Goldberg (1990).

Typically, personality questionnaires are lengthy and an excessive amount of time can be spent completing the measure, entering the data, and interpreting the results. Due to the labour involved in this process developers are often requested to reduce the length of questionnaires and by some means maximise the resulting information (Wang, Chen, & Cheng, 2004).

As mentioned by Tuerlinckx, Boeck, and Lens (2002) the accuracy of information provided by lengthy questionnaires comes into question for two main reasons: from the developer's perspective longer questionnaires tend to include lower quality items such as filler items, non-specific items, and items that are included solely to improve reliability; from the participant's perspective longer questionnaires increase the likelihood of losing concentration and making inaccurate responses through boredom, laziness, or unknowingly responding in a repetitive manner. Tuerlinckx et al. also found that questionnaire length significantly correlated with the final score on their measure. They suggested that IRT models could be fitted to personality checklists in a way that could identify a point where test fatigue influences the responses of the participant. They termed this the 'drop-out' point and this was explained as a consequence of loss of attention and loss of patience as participants responded without having fully read the question.

To alleviate the issues that arise out of lengthy questionnaires this research seeks to improve the quality and measurement precision of an existing personality questionnaire by reducing it to the core items that provide the best information about the participant.

The questionnaire that will be used in this research is derived from the freely available online resource at http://www.ipip.ori.org/ipip developed by Goldberg (1990). Many researchers have used this resource (e.g. Chernyshenko, Stark, Chan, Drasgow, & Williams, 2001b; Guenole & Chernyshenko, 2005) including the organisation that provided the data for this study.

The original developers of these items indicated that these are preliminary items as only rudimentary procedures were applied in developing the scales (Goldberg, Johnson, Eber, Hogan, Ashton, Cloninger, & Gough, 2006). Goldberg et al. (2006) suggested that an IRT analysis would identify the highest quality items from this item-pool, and subsequently invited other researchers to perform this task. The results of such an analysis would be applicable for anyone who uses the items from their website however a preliminary search through the 100 plus articles on their website showed no indication of this task being achieved.

Test development has traditionally been performed using *Classical Test Theory* (CTT). However as questionnaires are completed at the item level, it is logical that they should also be developed and interpreted at the item level (Fletcher & Hattie, 2005). This form of analysis cannot be achieved through CTT and therefore an alternative method is necessary.

Theories Underlying Personality Test Development

Classical Test Theory and Item Response Theory

Hambleton and Jones (1993) compare and discuss the two major theories underlying test construction and development, CTT and IRT. They state that models cannot perfectly represent the test data they are associated with, and therefore the question in relation to which theory to use should be based on which will help create a model that will best guide the measurement process. The model strength is dependent on the assumptions that must be met in order to use the relevant framework. Hambleton and

Jones (1993) state that CTT models are often weak as the assumptions are easily met whereas IRT models are stronger as the assumptions are harder to meet. For example, in IRT the assumption is made that the set of items grouped under one label must only measure that single trait or ability and therefore unidimensionality (discussed below) must be satisfied when applying this theory (Raju, Laffitte, & Byrne, 2002). Conversely, CTT only assumes that the structure of a model is consistent when tested with different samples.

The majority of test development is currently performed using CTT. This is due to two main factors. Firstly, IRT is a statistically complex procedure and software was not available that made the process simple to utilise (McKinley, 1989), however this has now changed. Secondly, any new theory must be thoroughly tested and refined before it is applied to real data (Zickar, 1998). IRT has now gone through this process and can thus be used in mainstream testing. IRT has made big impacts on quantitative psychology as the underlying statistical base of IRT along with the development of computer technology has meant that computerised adaptive testing can now be performed. This combination gives the precision of classical tests with the efficiency of advanced software that can select an item that will obtain the most useful information (Ozer & Reise, 1994; Zickar, 1998). The key differences between CTT and IRT will show why IRT is quickly increasing in popularity and use.

Classical Test Theory
In its basic form CTT utilises three core concepts: the observable *test score*, the unobservable *true score* and the unobservable *error score*. CTT provides the assumption that the average *error score* (for the population that completed the test) is zero and hence the *true score* is derived directly from the *test score*. As this assumption is based on the average response to a group of items, two aspects of the data are lost. The first is the ability to assess individual responses, as the output statistics are derived from group averages rather than independent items. The second is that the process of averaging constrains the usefulness of the outcome statistics as no feature of the process indicates that the outcome could be generalised outside the sample from which they were derived, thus making the output statistics 'sample dependent'. As stated by Hambleton and Jones (1993) "this dependency reduces their utility".

Error is not estimated through the CTT procedure. This means that apart from the underlying construct, any other factor that may influence the participant's response is unaccounted for. In contrast, Gefen (2003) explains that every variable in a test introduces an element of measurement error that does not relate to the actual underlying construct. Some CTT models improve upon this basic assumption by indicating that there is measurement error but that the distribution of the error can be estimated using a predetermined curve, such as normal distribution (Hambleton & Jones, 1993). This addition improves the output statistics by identifying the error but does not give a true indication of the error associated with an item.

Item Response Theory

The issues that have been raised in regards to CTT (often produces weak models; loss of item information; sample dependency; unaccounted for error) are overcome by using IRT. IRT is a statistical theory about an individual's response to an item and how that relates to the relevant ability, trait, or construct that is being measured. There are two typical underlying assumptions involved in creating models within the IRT framework. The first pertains to the dimensional structure of the test data (Hambleton & Jones, 1993). This assumes that items that are grouped together are measuring one facet or category of information. This is referred to as *unidimensionality* as each item should only measure one unique factor (McKinley, 1989). The second relates to the form of the graph that represents the item. This graph is created using the data from the item (how people have responded) and a relevant IRT formula (Hambleton & Jones, 1993). The assumption in regards to unidimensionality will now be explained.

Unidimensionality

Gefen (2003) states that every item should only have one underlying construct. This means that items should only reflect their associated construct without significantly reflecting any other. This concept can be clarified through making the distinction between common variance and non-common variance. If two items in a test are hypothesised to measure the same construct then a proportion of the variance they capture is effectively in common. However, items generally do not have perfect measurement properties and therefore also capture other variance that is referred to as non-common variance. An item is not unidimensional when its non-common variance is

highly correlated with the non-common variance of another item, thus indicating that the items are capturing the variance of more than one dimension.

Although this analysis is important for assessing the strength of a model, the literature regarding unidimensionality is controversial. As Hattie (1984, 1985) describes, most indices of unidimensionality have some form of problem. Therefore great care should be taken when selecting which method of analysis is used (for a comprehensive review see Hattie, 1984, 1985). Despite these issues, it is important to assess unidimensionality and this is often performed through *Confirmatory Factor Analysis* (Gefen, 2003).

Item Characteristic Curve
The second IRT assumption pertains to the shape of the graph produced by each individual item. This graph or more specifically the line that is formed by the data on the graph is called the *Item Characteristic Curve.*

The *Item Characteristic Curve* (ICC) is a graphical representation of how and where an item works. The graph plots the probability of a correct response or endorsement of an answer, against ability or endorsement or a trait (McKinley, 1989). The principal of having an ability score is a fundamental difference to CTT that utilises test scores. That is because a person's ability is independent of (1) *the test they are completing*, (2) *the others that complete the test* and (3) *the other items in the test* (Hambleton & Jones, 1993). An example of this is that a person will have a lower score on a difficult test than they will on a simple test, however their ability will remain constant over both tests. Their ability should also remain constant over any other tests that measure the same construct, if completed at the same time. This signifies that ability (or endorsement) can be plotted on a continuum, and this continuum is dependent on the item itself and not the people who responded to the item. This gives all parameters estimated through IRT the property of invariance (McKinley, 1989) and hence the item parameters do not vary when used with different samples.

Sample Independence

McKinley (1989) states, "Item statistics that are obtained from the application of IRT models are independent of the sample of examinees to which a test (or other instrument) is administered". This is in contrast to traditional statistics where scores are stated as a percentage of correct responses and where the statistic most frequently used for comparison is the mean score. This traditional procedure indicates that the output statistics are only relevant to their sample of origin or a sample that has been shown to be very similar. Therefore, in order to obtain comparisons for people completing tests, organisations expend great effort building databases of different sample groups. Conversely, a single analysis can be performed through IRT and all respondents can be assessed on the same scale. In this way IRT avoids sample dependency and adds utility (McKinley, 1989).

IRT Models

There are two major families of IRT models, dichotomous and polytomous. Dichotomous models are for items that have binary answers: *yes* or *no*, *agree* or *disagree*, *1* or *2*. Polytomous models are for items with more than two responses (Ostini & Nering, 2006). Whether dichotomous or polytomous, all IRT models effectively include three estimation parameters: an item discrimination parameter 'a', a difficulty parameter 'b', and a guessing parameter 'c'. In the one-parameter model (or Rasch model) the 'a' is set at 1, 'c' is effectively set at 0, and the IRT formula estimates the 'b' parameter in order to produce the item graphs. In the two-parameter model (or logistic function) both the 'a' and the 'b' parameters are estimated by the formula. In the three-parameter model all three parameters are estimated (Baker, 2001; Hambleton, Swaminathan & Rogers, 1991).

IRT gives a true understanding of how an individual item operates through the use of item parameters. The discrimination or 'a' parameter is labelled as such because it illustrates how well an item differentiates between individuals, as an item with a high 'a' discriminates more than an item with a low 'a'. The item difficulty or 'b' parameter is labelled as such as the item graphs visually illustrate where on the continuum an item operates. Therefore, in regards to ability the value of the 'b' parameter will indicate whether the item operates in the low end of the scale, hence is an 'easy' item, or the

high end, hence is a 'difficult' item. The '*b*' parameter is also referred to as the 'response option location parameter' (Fletcher & Hattie, 2004) as the graph informs the user where the item best differentiates between individuals, i.e. between people at the low end or high end. It is of benefit to the user to have items in a scale that operate in different areas of the personality continuum. This means more of the information about the latent variable is captured and therefore it can be better understood and is more practical.

An important difference between dichotomous and polytomous models is in regard to the amount of '*b*' parameters that are estimated. In dichotomous models '*b*' represents the threshold point between a respondent choosing category 1 or category 2, e.g. 1 = '*yes*', 2 = '*no*'. However, polytomous models require the estimation of additional '*b*' parameters due to the multiple response options. A polytomous model with, for example, five categories would have four '*b*' parameters labelled '*b1*', '*b2*', '*b3*' and '*b4*', each representing the threshold point between the five category options. A further difference between these models is that the main item graph for dichotomous items is the *Item Characteristic Curve*, whereas for polytomous items this is referred to as the *Category Characteristic Curve* (Fletcher & Hattie, 2004).

There are three main models available when using polytomous IRT. Two of these are *Rasch* type models (one parameter models), namely the *Partial Credit* (PC) model and the *Rating Scale* (RS) model. These only estimate one parameter due to the "Principle of specific objectivity" (Ostini & Nering, 2006), which is derived from the theory that person parameters (which influence the item discrimination parameter) should be separate from the item parameters. Therefore the '*a*' parameter remains constant and only the '*b*' parameters are estimated by the formula (Ostini & Nering, 2006). The PC model assumes that responses are ordered meaning that as a respondent successfully progresses through the items their ability level also increases (Fletcher & Hattie, 2004). The name of this model is due to the fact that a correct response to the first part of an item and not the second part still receives partial credit. The RS model is similar to the PC model and is derived from the same underlying principles (Ostini & Nering, 2006). The third option is the *Graded Response* (GR) model (Samejima, 1969), which does not assume that item discrimination is the same between items.

Theoretically, as the *Rasch* models focus on correct or incorrect responses they are not well suited to personality testing in comparison to the GR model, which is more useful for trait endorsement data. This is illustrated through many studies that have selected this model for the development of personality questionnaires (Bolt, Hare, Vitale, & Newman, 2004; Fletcher & Hattie, 2004; Gomez, Cooper, & Gomez, 2005). For a complete description of the GR model refer to *'Polytomous Item Response Theory Models'* by Ostini and Nering (2006). In addition, a comparison of the application of different IRT models to personality data can be seen in Chernyshenko et al. (2001b).

Typical Methods of Questionnaire Development

Two key aspects of questionnaire development are in regard to (1) *the way in which items fit together in a factor* and (2) *the way factors fit together in a model*. The first of these aspects, item to factor fit, is typically measured through reliability analyses. Churchill (1979) stated that reliability should be the first measure calculated to assess the quality of a factor, the most common measure of which is the *Cronbach's Alpha*. Higher reliability is achieved by having items that load well together. This may signify that the items are asking the same question in a different way. For this reason item to factor fit and also factor to model fit are better measured through unidimensionality analyses such as can be performed through CFA.

CFA has been used for many studies assessing the fit of models for personality inventories (Borkenau & Ostendorf, 1990; Raju et al., 2002; Guenole & Chernyshenko, 2005). In these studies CFA has been stated as an appropriate methodology for confirming the underlying structure of an inventory. An important aspect of these analyses is that they are performed not only to confirm the hypothesised structure, but also to reject other plausible models. Additionally, CFA provides the means to test for unidimensionality, which is of critical importance for test validity (Gefen, 2003). If unidimensionality is not satisfied this can lead to incorrect interpretations of the strength of relationships within the model (Chernyshenko, Stark, & Chan, 2001a). The primary concern addressed through CFA in personality literature is the factor structure of each questionnaire, as there are many opinions regarding which factor structure best describes personality data.

Factor structure disagreement has been a major catalyst for the different forms of personality questionnaires currently available. This conflict is mainly caused by the difference of opinion in regard to what is actually being measured (Eysenck, 1992). Eysenck is the primary personality theorist opposing the FFM and alternatively proposes a three-factor model using *Extraversion, Neuroticism* and *Psychotocism* (Guenole & Chernyshenko, 2005). Ones and Viswesvaran (1998) propose a two-factor model where *Conscientiousness, Agreeableness, Emotional stability* load on one factor, and *Extraversion* and *Intellect* on the second. The 16PF (Conn & Rieke, 1994) is a FFM however the emphasis is on the 16 lower-order factors rather than the five higher-order factors (Chernyshenko et al., 2001a). In each of these cases the factor structure is proposed based on developer preference.

Researchers who question the validity of the design of other measures often test the proposed factor structures with their own data. Chernyshenko et al. (2001a) state that although the 16PF is the most influential and well-researched self-report personality inventory developed in the past 50 years, there was still a need for the unidimensionality of the 16 non-cognitive scales in the 16PF and the hierarchical factor structure of the inventory to be investigated. This was motivated by the recent development of the test from the fourth to the fifth edition as many of the items had considerably changed. Some had minor changes (such as subtle rewording) and many had been discarded and replaced with items that were completely new to the measure. Only 22% of the 185 items in the measure were exactly as they were in the fourth edition, therefore it was of determined that the factor structure should be reconfirmed. Their analysis using a hierarchical *Exploratory Factor Analysis* resulted in a confirmation of the hypothesised factor structure as the 185 items loaded on 16 first-order factors, which loaded on five second-order factors.

Being in its fifth addition the 16PF is an example of a personality questionnaire that is subject to continuous development and improvement (Gerbing & Tuley, 1991). The item level development of this test means that item properties are theoretically constantly being improved with the additional data providing means for the ongoing analysis. Many questionnaires go through the development process (Costa & McCrae, 1997) as this improvement is of empirical benefit to the end users.

Goldberg's Online Inventory

Another example of an ongoing test development process is seen through the constantly updated public domain instrument developed by Goldberg (1990) available at http://www.ipip.ori.org/ipip. Goldberg has made over 2000 items available for researchers, teachers, students, small organisations, or any person who would like to make use of this item bank. Many of the items are based on the major personality inventories that have been mentioned in this study. The items have been correlated with the original scales, redundant items were discarded based on similar wording to other items, reliability analyses were performed and the items have been categorised for those who wish to use them (a full description of this process is available in Goldberg et al., 2006). This has meant that researchers from around the world can use this resource without cost, so they can confirm or reject their personality research hypotheses. As stated, Goldberg has invited any researchers to develop these items using applications such as IRT in order to improve the quality of these scales.

IRT Research

Current personality research has shown some movement towards analysis with IRT. This is a statistically complex procedure (McKinley, 1989) however the detailed information that is provided is invaluable for those who see the importance of measurement precision.

Fletcher and Hattie (2004) applied IRT to a 70-item *Physical Self-Description Questionnaire* (PSDQ) and identified good items, mediocre items that should be reworded, and poor items that should be discarded due to the limited amount of unique information they provided. Through this process Fletcher and Hattie (2004) showed how to minimise the length of the questionnaire by identifying items that captured the best quality information. This item level analysis is only available through IRT.

A further application of IRT is seen through the development of the *Asian Values Scale* through to the *Asian Values Scale- Revised* (Kim & Hong, 2004). In this analysis it was stated that the original 35-item scale was developed using CTT through reliability and validity analyses. The scale was revised using IRT in an attempt to improve the measurement properties of the scale. Their analysis through the use of the *Rasch Model*

resulted in a reduction from their original list down to 25 items and a reduction of response options from the original 7-point Likert-scale down to a 4-point Likert-scale. Hong, Kim and Wolfe (2005) performed a similar IRT analysis with the use of a *European American Values Scale* (EUVS). In this study the EUVS had 18 items, which had been revised from an original list of 180 items. This original list was then subjected to the IRT analysis and 25 items were selected along with the same reduction of 7 response options down to 4 response options for the EUVS-*Revised*. The results of these two studies stemmed from the valuable item level information that was gained through the use of the IRT graphs. It is also interesting to note that in regards to the Likert-scale both of these personality analyses were reduced from 7 options down to 4 options.

Although Kim and Hong (2004) and Hong et al. (2005) opted for a scale wide response option reduction, this is not always the case. Through the IRT analysis performed by Fletcher and Hattie (2004), no changes were made to the questionnaire however recommendations were given. These included items that should be kept as the core of a future revised questionnaire, items that suited the current Likert-scale, items that would be better suited to a dichotomous scale, and items that needed rewording and retesting in order to be included in the revised questionnaire. Fletcher and Hattie (2004) utilised Samejima's GR model, which estimates all three parameters involved in polytomous IRT, whereas Kim and Hong (2004) and Hong et al. (2005) selected the one-parameter Rasch model. Better quality information is typically gained through using the three-parameter model over the one-parameter model, however a larger sample size is needed (Tabachnick & Fidell, 2007) which can limit the model selection.

Gomez et al. (2005) also selected Samejima's GR model for their analysis of two behaviour-based scales. Rather than focus on individual items as was shown through the studies mentioned above, Gomez et al. assessed the information captured by the whole scale. They found that the items were generally good however they only provided information about their latent traits from the moderately low to the moderately high areas of the continuum thus signifying issues for the psychometric properties of the scales. Recommendations were made for additional items to capture information at each end of the continuum.

IRT was also used in a psychometric analysis performed by Tuerlinckx et al. (2002). An interesting component of this analysis was the decision to split their dataset between males and females and use this as a form of cross-validation. From this procedure they were able to illustrate similar findings between the two separate groups and conclude that the findings from one part of their study cross-validated the findings from another part. In regards to any questionnaire development, the process of cross-validation with different samples is highly recommended (Tabachnick & Fidell, 2007).

Objectives and Hypotheses for the Current Study

Longer questionnaires often include redundant items that can decrease the measurement precision of the test (Tuerlinckx et al., 2002). Through IRT the best quality items in a questionnaire can be identified. Therefore a model produced with items selected through IRT should show much better fit than a model that includes redundant items in terms of both unused response options and items that capture little information (Fletcher & Hattie, 2005). IRT assumes unidimensionality and therefore any factors analysed should be assessed using this principle (Raju et al., 2002). For this reason, this study will perform a test of model-fit on the original questionnaire using CFA (*Model 1*), followed by the deletion of poor items as shown through these analyses, after which another CFA (*Model 2*) will be run in order to measure the improved fit of the model. This will be the first stage of analysis and it is hypothesised that the fit of the model will improve.

In order to further identify and select the best items for each factor, IRT analyses will be performed (see Gomez et al., 2005; Fletcher & Hattie, 2004; Kim & Hong, 2004; Hong et al., 2005; and Chernyshenko et al., 2001b). Tabachnick and Fidell (2007) stated that a minimum of three variables should be used to measure a factor. Accordingly, the three best items will be identified for each lower-order factor and these will be combined for a final reconstituted CFA model (*Model 3*). This will be used for comparisons with the previous two models. It is hypothesised that the model fit will once again significantly improve from this procedure.

This reconstituted questionnaire will show far greater measurement precision than its original state with additional efficiency of use due to its reduced length and its lack of redundant items. As information is additive, the deletion of redundant items will lead to

lower overall information, however due to the selection of high-quality and high-information items it is expected that the information reduction will be minimal in comparison to the item reduction. Furthermore, the results of this identification process will be of great value to any users of Goldberg's online resource due to the fact that the parameter estimates are dependent on the individual items and not the sample that was used in this analysis hence the resulting item selections can be freely generalised and are thus highly relevant to many individuals and organisations.

Method

Participants

This study used data provided by an Organisational Psychology consultancy in New Zealand. The sample consisted of 973 adults, 376 of which were female and 597 of which were male. Participants were aged between 16 and 80 ($M = 42.40$, $SD = 8.93$). The majority of participants described their ethnic/cultural background as NZ European ($n = 774$), followed by Other European ($n = 102$), Maori ($n = 80$), Asian ($n = 22$), Pacific Islander ($n = 16$), and Other Ethnic Group ($n = 11$).

Measure

Original State of Questionnaire

This personality questionnaire is hypothesised to be a 3-stage higher-order model. The items in the questionnaire were derived from the online resource developed by Goldberg (1990) and are modelled in the design of the *Big-Five*. Therefore, this questionnaire includes five factors that give an indication of personality and these are labelled as follows: *Extraversion and Impact, Emotional Management, Intellectual Preferences, Interpersonal Style,* and *Self Management and Drive.* The relationship between the traditional *Big-Five* factors and the factors in this questionnaire is shown in Fig. 1.
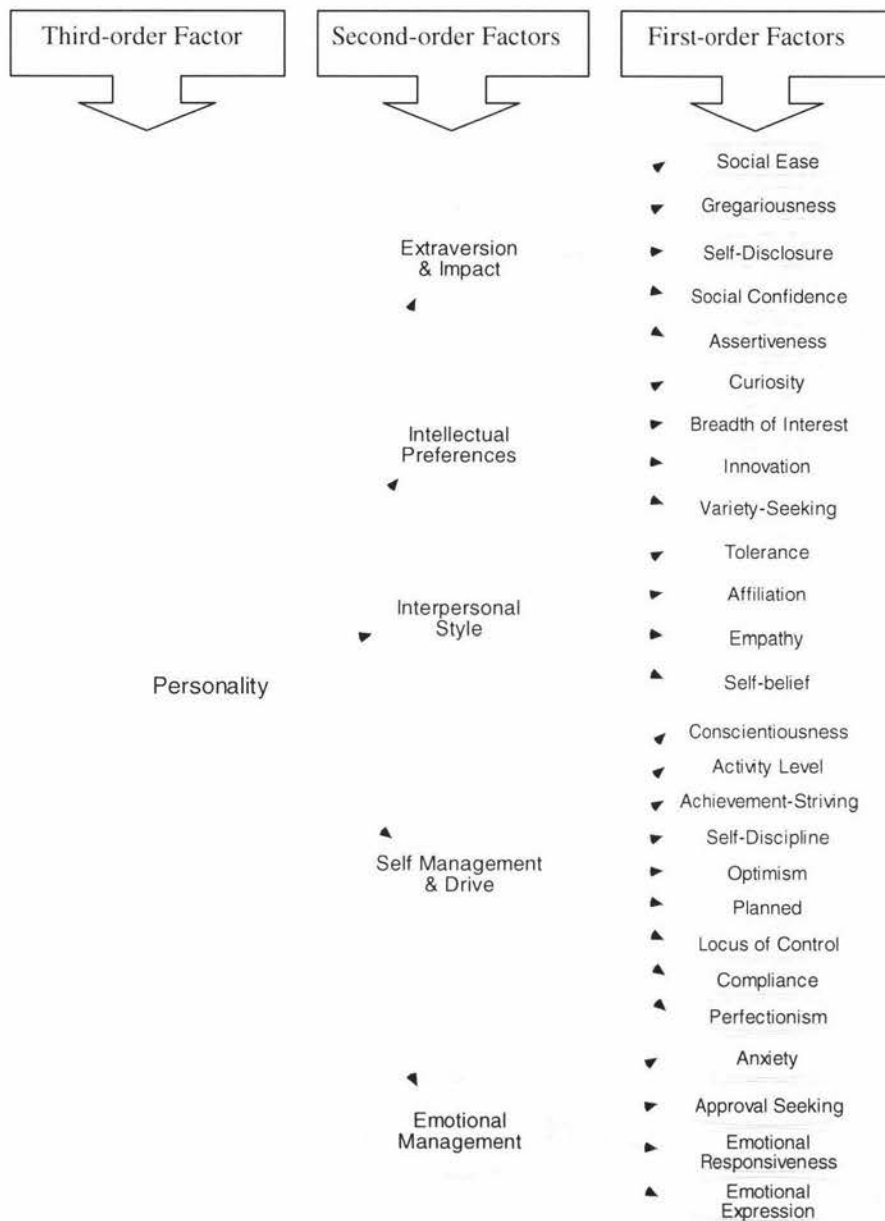
Fig. 1

Relationship between Traditional Big-Five Factors (left) and the Five Second-order Personality Factors (right) in this Questionnaire

A model depicting the hypothesised structure of this personality questionnaire is shown in Fig. 2 with 26 first-order factors, five second-order factors and a single third-order factor. Chernyshenko et al. (2001b) also used this terminology to describe the factors in their analysis of the 16PF.

Fig. 2

Factor Structure of Questionnaire

Third-order Factor	Second-order Factors	First-order Factors

Extraversion & Impact
- Social Ease
- Gregariousness
- Self-Disclosure
- Social Confidence
- Assertiveness

Intellectual Preferences
- Curiosity
- Breadth of Interest
- Innovation
- Variety-Seeking

Interpersonal Style
- Tolerance
- Affiliation
- Empathy
- Self-belief

Personality

Self Management & Drive
- Conscientiousness
- Activity Level
- Achievement-Striving
- Self-Discipline
- Optimism
- Planned
- Locus of Control
- Compliance
- Perfectionism

Emotional Management
- Anxiety
- Approval Seeking
- Emotional Responsiveness
- Emotional Expression

There were 246 individual items associated with 26 first-order factors in this questionnaire. These were divided between five second-order factors. The amount of items affiliated to each first-order factor is shown in Table 1 and is listed in Appendix 1.

Many of the items in this list cross-loaded to one or more factors, therefore the total number of items shown in Table 1 (268) is greater than the total number of individual items (246). The original questionnaire included a lie scale (17 items) in order to identify participants who did not respond truthfully. As this was not directly associated with the personality factors, it was analysed separately from the rest of the items and is shown at the end of the appendices (Appendix 6).

Table 1

Amount of Items in Each First-order Factor

Second-order Factor	First-order Factor	Items
Extraversion & Impact		
	Social Ease	10
	Gregariousnes	10
	Self-Disclosure	10
	Social-Confidence	10
	Assertiveness	10
Intellectual Preferences		
	Curiosity	10
	Breadth of Interest	10
	Innovation	10
	Variety-Seeking	10
Interpersonal Style		
	Tolerance	11
	Affiliation	10
	Empathy	8
	Self-belief	10
Self-Management & Drive		
	Conscientiousness	10
	Activity Level	10
	Achievement-Striving	10
	Self-Discipline	10
	Optimism	10
	Planned	10
	Locus of Control	20
	Compliance	10
	Perfectionism	9
Emotional Management		
	Anxiety	10
	Approval Seeking	10
	Emotional Responsiveness	10
	Emotional Expression	10
	Total items	268

For all of the items in this questionnaire participants were asked to answer how accurately the item described them using a 5-point Likert-scale:'1'- *Very Inaccurate*, '2'- *Moderately Inaccurate*, '3'- *Neither Inaccurate nor Accurate*, '4'- *Moderately Accurate*, and '5'- *Very Accurate*. As many items were negatively worded (126 out of 268) these were recoded into the same direction as the positively worded items.

Procedure

Split of Dataset

To enhance the validity claims for this measure, the data were randomly split into two files so that the factor structure of the measure could be tested with the full model using the first data set of 484 participants and then with a reduced length scale using the second data set of 489 participants. Cross-validation is a typical procedure used to increase the strength of statistical analyses. For a good example see Tuerlinckx et al. (2002). The full data set of 973 participants was used for the IRT analysis.

Cross-loaded Items

When items represent more than one construct the interpretation of what they represent is difficult to discern. For factor integrity and interpretation items should only load on one factor. This personality questionnaire had 21 items that were suggested to measure more than one factor (Appendix 2). At the beginning of this study a decision was made to discard these items so that the CFA could be run and so that the principle of unidimensionality could be satisfied. The data for these cross-loaded items was not deleted so that they could be reanalysed if any of the first-order factors failed to converge in the initial analysis. Discarding the 21 cross-loaded items left 225 items for *Model 1*. The fit statistics of this model served as a base line for comparisons.

Confirmatory Factor Analysis

To assess the degree of unidimensionality a model was specified for each individual first-order factor resulting in a total of 26 CFAs. Subsequently, the first-order factors were combined with their associated second-order factor in order to create five second-order CFAs. These were then combined with a higher-order *Personality* factor to create the total model that was used for the comparisons (see Fig. 2 above). Three total model CFAs were calculated to illustrate each stage of the development and selection process.

Model 1: The first model is referred to as *Model 1* (225 items) and includes the original length first-order factors after the cross-loaded items were discarded.

Model 2: The *Model 1* first-order factors were then assessed for model fit. Two of the 26 first-order factors failed to converge. The cross-loaded items were added back to

these two scales and they were reanalysed and successfully converged (this process will be explained in the *CFA Results* section). In order to satisfy the requirements of unidimensionality, poor items were deleted from all 26 first-order factors in *Model 1* based on the *Squared Multiple Correlation*. The remaining items from these first-order factors were then reformed into a model referred to as *Model 2* (187 items).

Model 3: The items from *Model 2* were then subjected to the IRT analysis. Three items from each of the 26 first-order factors were selected and combined into a final model referred to as *Model 3* (78 items).

All CFA models were calculated using AMOS 4.0 (Arbuckle, 1999). When an error term was reported in the model to have negative variance the error-variance of the specific parameter was fixed to .001 as is acceptable under these circumstances (Byrne, 2001). Error-variance was fixed to .001 twice for *Model 1*, once for *Model 2*, and twice for *Model 3*.

Model Fit

Fit indices typically reported in *Confirmatory Factor Analyses* are the *Goodness of Fit Index* (GFI: Tanaka & Huba, 1984), the *Tucker-Lewis Index* (TLI: Bollen, 1989) and the *Comparative Fit Index* (CFI: Bentler, 1990), where > .90 indicates adequate model fit for each of these three fit indices. One further fit statistic referred to as one of the best model fit indicators (Fletcher & Hattie, 2004) is the *Root Mean Square Error of Approximation* (RMSEA: Steiger & Lind, 1980), where $.00 < .05$ indicates close fit, $> .05 < .08$ indicates reasonable fit, $> .08 < .10$ indicates tolerable fit, and $> .10$ indicates poor fit (Browne & Cudeck, 1993). Examination of these fit statistics indicates whether or not a reasonable fit of the data to the model has been achieved.

Item Response Theory

IRT was then used to identify the best three items for each of the 26 first-order factors in *Model 2*. This was achieved using the polytomous GR model (Samejima, 1969). The items selected from the IRT analysis formed *Model 3*.

The Graded Response Model

The GR model was used to produce many different informative graphs and item statistics. These graphs illustrate a wealth of item level information that is not available with traditional statistical analyses (Fletcher & Hattie, 2004). The graphs used in the results section of this study include *Category Characteristic Curves* (CCC), *Operating Characteristic Functions* (OCF), *Item Information Functions* (IIF), and both *Test* and *Scale Information Functions* (TIF & SIF). These graphs illustrate the amount of information captured by each item, first-order factor, second-order factor, and complete model.

Polytomous IRT Graphs

To create the polytomous IRT graphs the '*a*' and '*b*' parameters were extracted from the raw data using a programme developed by Thissen (1991) called MULTILOG 6.0 (BILOG (Mislevy & Bock, 1991) is for dichotomous data; MULTILOG is for polytomous data). Individual files were created in SPSS (Version 13) for each first-order factor using the complete set of data ($n = 973$). These were converted to files that could be used with MULTILOG so that the discrimination (*a*) and difficulty (*b1*, *b2*, *b3*, and *b4*) parameters for all of the items could be produced (see De Ayala (1993) for a more detailed description of this process). The '*a*' and '*b*' parameters were entered into a MICROSOFT EXCEL spreadsheet and Samejima's (1969) GR model formula was used to produce the graphs.

Method for Selection of Three Best Items for each First-order Factor

Items were selected based on the item properties illustrated in the item level graphs. These properties include:

    (1) *the shape of the graph,*

    (2) *the location of item information,*

    (3) *total item information,*

    (4) *the use of all the response options,* and

    (5) *the combination of the items in the first-order factor,* including the

        a.   *item information location* and the

        b.   *item wording*

(1) The Shape of the Graph

The *Operating Characteristic Function* (OCF) for item Q005 (Fig. 3) illustrates certain properties that made this a good item. The area under each individual curve is effectively the information that is captured by that response option and in the OCF for item Q005 each individual response option had a high peak that was separate from the other peaks, meaning that each option captured a significant amount of unique information.

Fig. 3

CCC, OCF, and IIF for Social Ease Item Q005

Category Characteristic Curve		Operating Characteristic Function	Item Information Function Curve
a = Item discrimination parameter		1- Very Inaccurate	
b1 = Intersection between option 1 & 2		2- Moderately Inaccurate	
b2 = Intersection between option 2 & 3		3- Neither Inaccurate nor Accurate	
b3 = Intersection between option 3 & 4		4- Moderately Accurate	
b4 = Intersection between option 4 & 5		5- Very Accurate	

Extraversion and Impact   *Social Ease*

Q005   *Make friends easily.*

a	b1	b2	b3	b4
2.14	-3.35	-1.97	-0.96	0.72

(2) Location of Item Information

Item information should be evenly dispersed between response options and should only operate in a small section of the personality continuum. This means an item can have a more accurate degree of differentiation between individuals on the trait being measured. As seen in the CCC for item Q005 (Fig. 3), this is an example of a good item whereas item Q142 (Fig. 4) is an example of a poor item.

Fig. 4

CCC, OCF, and IIF for Tolerance Item Q142

Interpersonal Style   ***Tolerance***

Q142   *Believe that others have good intentions.*

a	b1	b2	b3	b4
0.80	-12.39	-5.41	-2.60	1.88
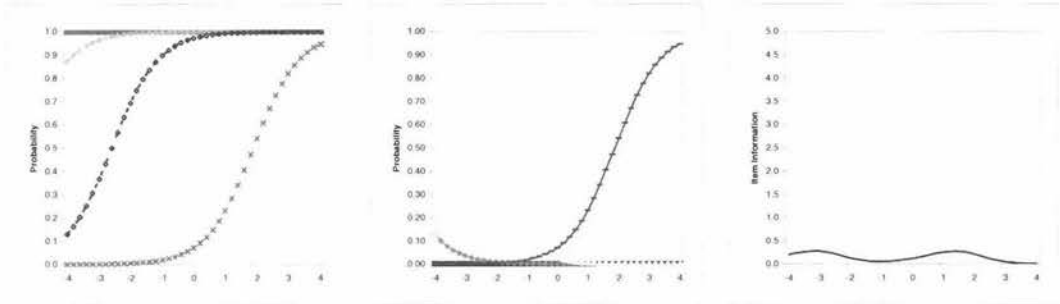
(3) Total Item Information

An additional aspect to notice when comparing items Q005 and Q142 is that the total information shown in the IIF was considerably lower in item Q142 from the '*Tolerance*' first-order factor. This also indicates that it is a poor item.

Two other items from the '*Tolerance*' first-order factor are shown in Fig. 5.

Fig. 5

CCC's, OCF's, and IIF's for Three of the First-order Factor Items for Tolerance

Interpersonal Style   ***Tolerance***

Q104   *Am a bad loser.*

a	b1	b2	b3	b4
0.75	-4.29	-1.88	-0.71	1.64
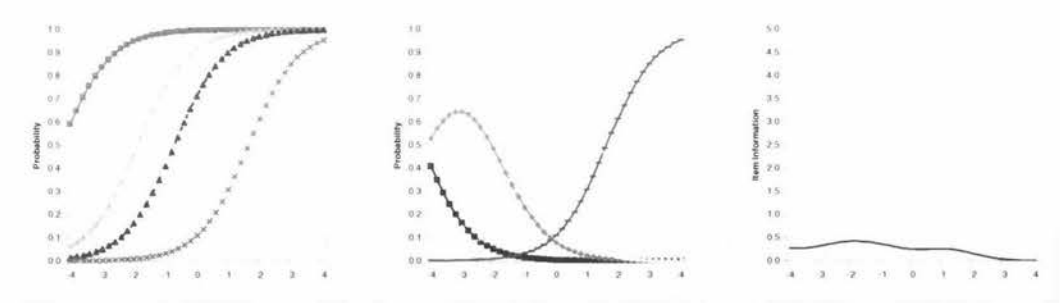
Interpersonal Style   ***Tolerance***

Q164   *Lay down the law to others.*

a	b1	b2	b3	b4
0.51	-8.14	-2.55	0.16	3.66

- 22 -
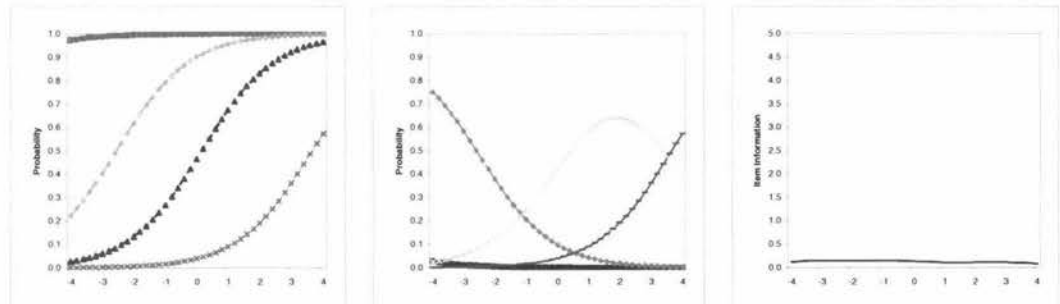
Bolt et al. (2004) suggest that 'a' parameters need to be over 1.00 to indicate reasonable discrimination whereas these two items each had 'a' parameters of 0.80 or less and hence captured very little information. Items that were shown to have information levels similar or worse than these were categorised as poor items.

(4) The Use of All the Response Options

Although Q094 (Fig. 6) has high peaks, high information, and good information location, the OCF shows that the information captured by response option '3' (*Neither Accurate nor Inaccurate*) was also captured by options '2' and '4'. This means that option '3' was effectively redundant and therefore this item was not suited to a 5-point Likert-scale questionnaire. Consequently, items such as this were not selected.

Fig. 6

CCC, OCF, and IIF for Innovation Item Q094

Intellectual Preferences   *Innovation*
Q094   *Can't come up with new ideas.*

a	b1	b2	b3	b4
2.09	-3.69	-2.41	-2.04	-0.5
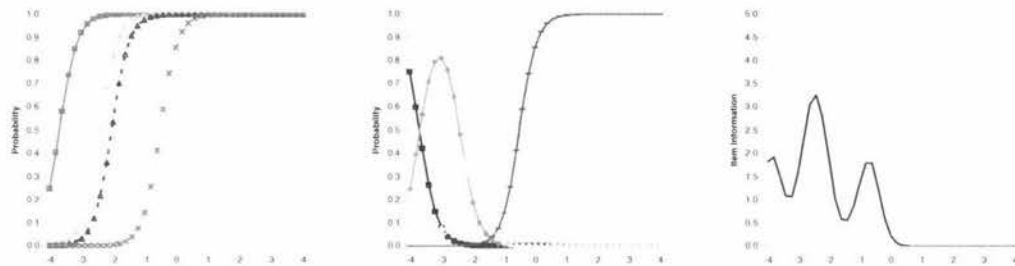
(5) The Combinations of Items in a First-order Factor

The item combinations were also analysed in order to maximise the variation amongst the wording of the items in each first-order factor and to capture information from different parts of the personality continuum. Examples are given below.

(5a) Information Location: Of the three items seen below in Fig. 7, two were selected for the final model. Although items Q021 and Q026 had higher 'a' parameters than Q032, the areas under the graphs of Q021 and Q026 (as can be seen in each IIF) were very similar, illustrating that they captured almost the same information. For this reason it was preferred to select only one of these items and then select a different item that captured different information, such as item Q032.

Fig. 7

CCC's, OCF's, and IIF's for Three of the Anxiety First-order Factors Items

Emotional Management   *Anxiety*

Q021   *Have frequent mood swings.*

a	b1	b2	b3	b4
1.66	0	1.25	2.19	3.78

Emotional Management   *Anxiety*

Q026   *Get upset easily.*

a	b1	b2	b3	b4
2.01	-0.59	0.97	1.81	3.51

Emotional Management   *Anxiety*

Q032   *Am not easily bothered by things.*

a	b1	b2	b3	b4
1.33	-1.68	0.48	1.38	3.66
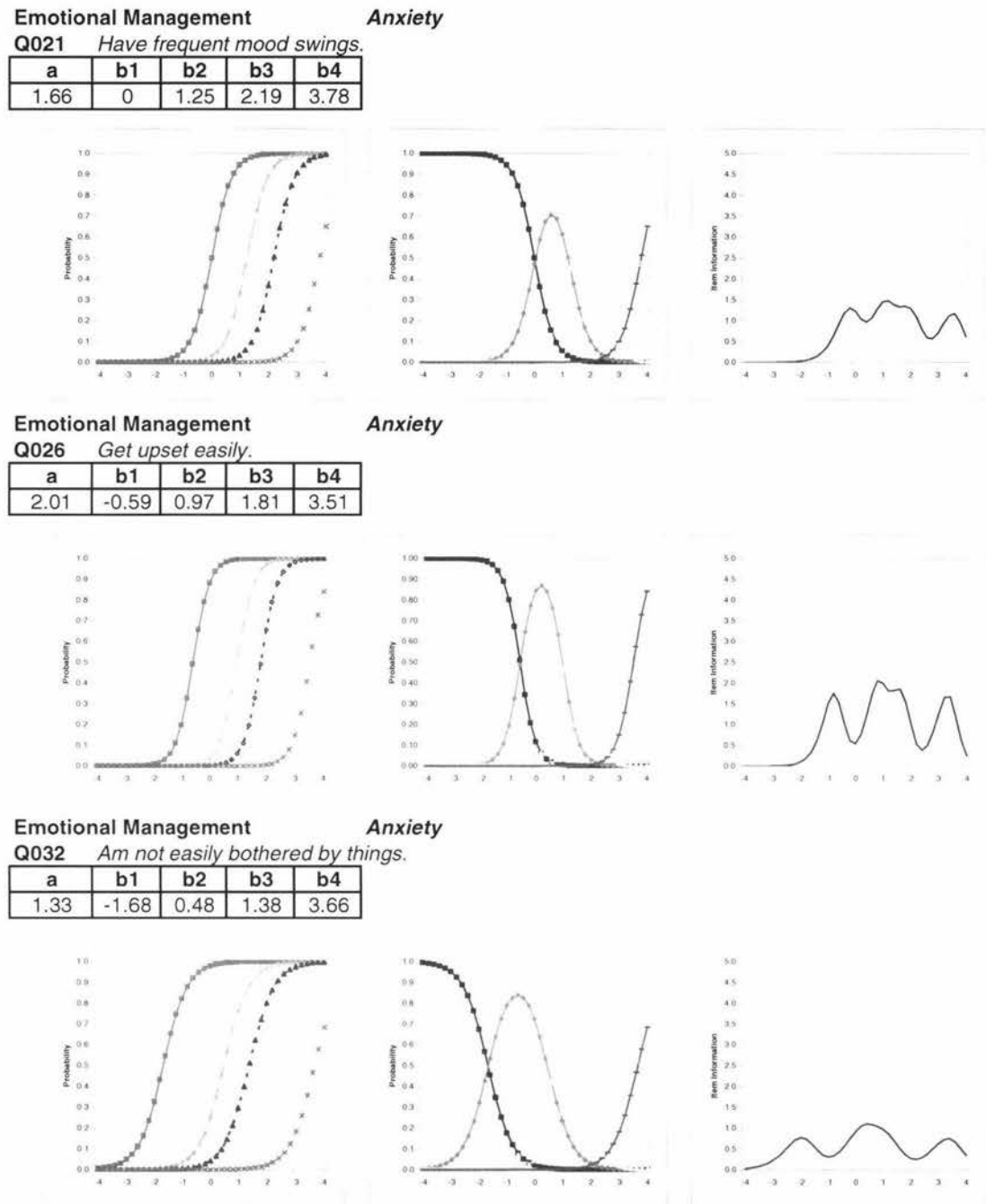
(5b) Item Wording: The wording of an item also provides insight in regards to which items to select. For the first-order factor '*Empathy*', two of the items were worded as follows: Q116- '*Make people feel welcome*' and Q236- '*Take time out for others*'. The third choice was between two options: Q280- 'Am concerned about others' and Q099- '*Reassure others*'. The highest '*a*' parameter belonged to Q099, however as this item referred to a behaviour, as did the first two, it was rejected in favour of Q280 which refers to an emotion and hence was semantically different.

Results

CFA Results

Each first-order factor from *Model 1* was individually assessed for unidimensionality using AMOS. The fit statistics from these 26 first-order factors are shown in Table 2.

Table 2

Fit Statistics for all 26 First-order Factors

Second-order Factor	First-order Factor	Items	Alpha	χ^2	DF	Prob	GFI	TLI	CFI	RMSEA
Extraversion and Impact (40 items)										
	Social Ease	7	0.80	35.60	14	0.00	0.98	0.96	0.97	0.06
	Gregariousness	10	0.85	267.37	35	0.00	0.89	0.81	0.85	0.12
	Self-Disclosure	9	0.83	446.33	27	0.00	0.85	0.65	0.74	0.18
	Social-Confidence	6	0.75	132.44	9	0.00	0.92	0.69	0.81	0.17
	Assertiveness	8	0.79	128.72	20	0.00	0.93	0.84	0.89	0.11
Intellectual Preferences (32 items)										
	Curiosity	8	0.83	184.51	20	0.00	0.91	0.81	0.86	0.13
	Breadth of Interest	7	0.69	185.79	14	0.00	0.89	0.62	0.74	0.16
	Innovation	9	0.80	192.67	27	0.00	0.91	0.80	0.85	0.11
	Variety-Seeking	8	0.76	100.97	20	0.00	0.94	0.86	0.90	0.09
Interpersonal Style (38 items)										
	Tolerance	11	0.69	106.96	44	0.00	0.96	0.87	0.89	0.05
	Affiliation	10	0.59	215.88	35	0.00	0.91	0.55	0.65	0.10
	Empathy	7	0.75	46.06	14	0.00	0.97	0.93	0.95	0.07
	Self-belief	10	0.78	405.45	35	0.00	0.82	0.60	0.69	0.15
Self Management and Drive (78 items)										
	Conscientiousness*	2	0.36	Unidentified			1.00		1.00	0.25
	Activity Level	10	0.71	293.02	35	0.00	0.88	0.57	0.67	0.12
	Achievement-Striving*	9	0.73	20.76	27	0.80	0.99	1.01	1.00	0.00
	Self-Discipline	5	0.79	89.41	5	0.00	0.93	0.75	0.88	0.19
	Optimism	8	0.66	166.30	20	0.00	0.91	0.65	0.75	0.12
	Planned	8	0.72	88.09	20	0.00	0.96	0.85	0.89	0.08
	Locus of Control	17	0.80	354.20	119	0.00	0.91	0.83	0.85	0.06
	Compliance	10	0.68	256.14	35	0.00	0.89	0.57	0.66	0.11
	Perfectionism	9	0.76	216.73	27	0.00	0.91	0.73	0.80	0.12
Emotional Management (37 items)										
	Anxiety	10	0.83	144.72	35	0.00	0.94	0.89	0.92	0.08
	Approval Seeking	10	0.65	216.46	35	0.00	0.91	0.63	0.71	0.10
	Emotional Responsiveness	9	0.86	142.68	27	0.00	0.94	0.90	0.93	0.09
	Emotional Expression	8	0.64	150.15	20	0.00	0.92	0.63	0.74	0.12
Total items		**225**								

* First-order Factors that failed to converge

Second-order Factor CFA for Model 1

A second-order CFA was then prepared for each second-order factor with its associated first-order factors. An example of one of these CFAs (for *Extraversion and Impact*) is shown along with the fit statistics in Fig. 8 with the remaining second-order factors shown in Appendix 3.

Fig. 8

CFA for Second-order Factor Extraversion and Impact as Part of Model 1

Items	χ^2	DF	Prob	GFI	TLI	CFI	RMSEA
40	2886.44	735	0.00	0.74	0.70	0.72	0.08
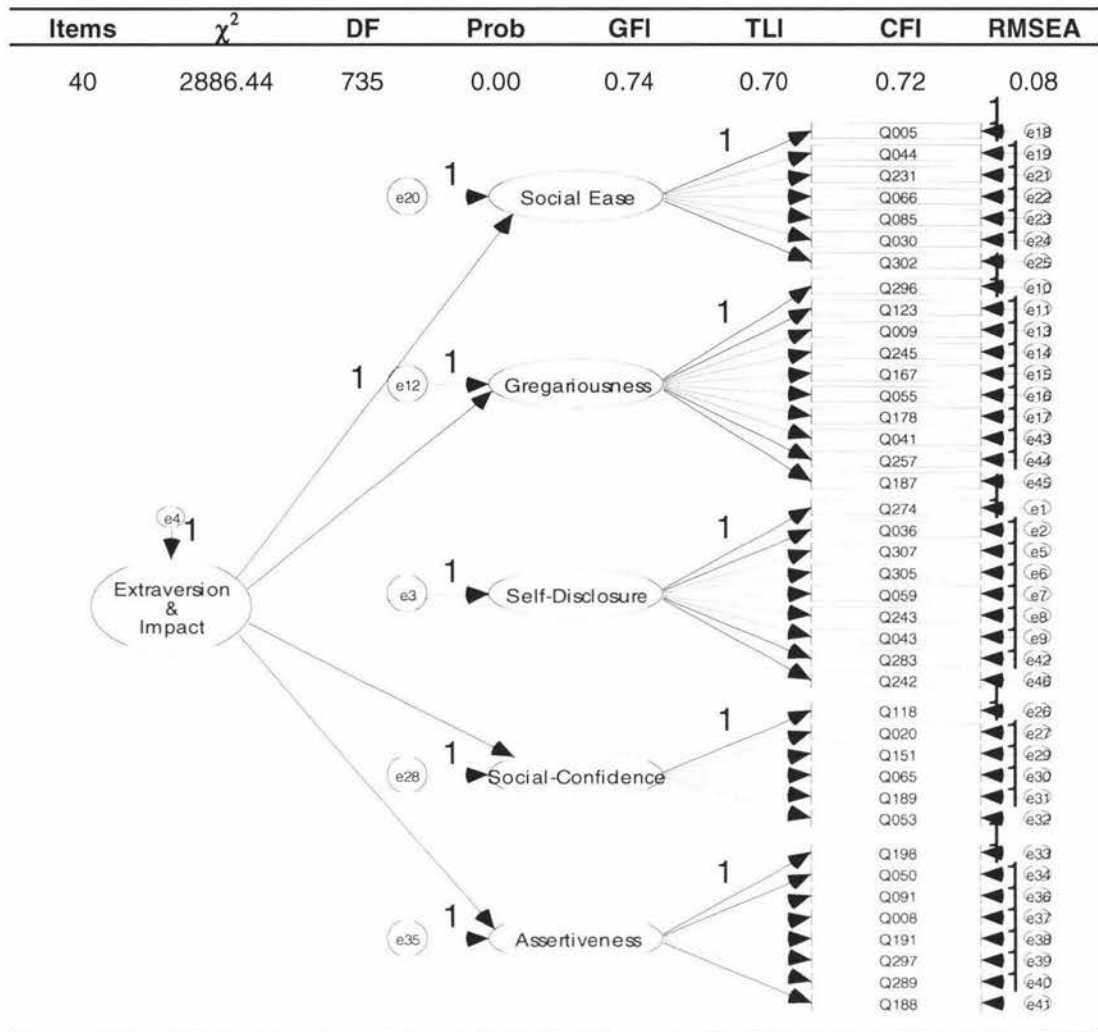
The fit statistics for the five second-order factors are shown in Table 3.

Table 3

Fit Statistics for the CFA of each Second-order Factor ready for Model 1

Second-order Factor	Items	χ^2	DF	Prob	GFI	TLI	CFI	RMSEA
Extraversion and Impact								
	40	2886.44	735	0.00	0.74	0.70	0.72	0.08
Intellectual Preferences								
	32	1748.09	460	0.00	0.79	0.73	0.75	0.08
Interpersonal Style								
	38	2085.83	662	0.00	0.79	0.64	0.66	0.07
Self Management and Drive								
	78	6689.77	2916	0.00	0.70	0.64	0.65	0.05
Emotional Management								
	37	1905.73	626	0.00	0.79	0.73	0.75	0.07
Total	**225**							

The RMSEA was at a reasonable level for the five CFAs shown in Table 3. In all cases the GFI, CFI, and TLI are lower than the preferred minimum level (.90). The combining of the five second-order factors into a model with the inclusion of the third-order factor *Personality* resulted in a good level for the RMSEA, however the GFI, CFI, and TLI dropped considerably as is shown in Fig. 9.

Fig. 9

CFA for Model 1

Model	Items	χ^2	DF	Prob	GFI	TLI	CFI	RMSEA
Model 1	225	49436.32	24948	0.00	0.47	0.48	0.48	0.05
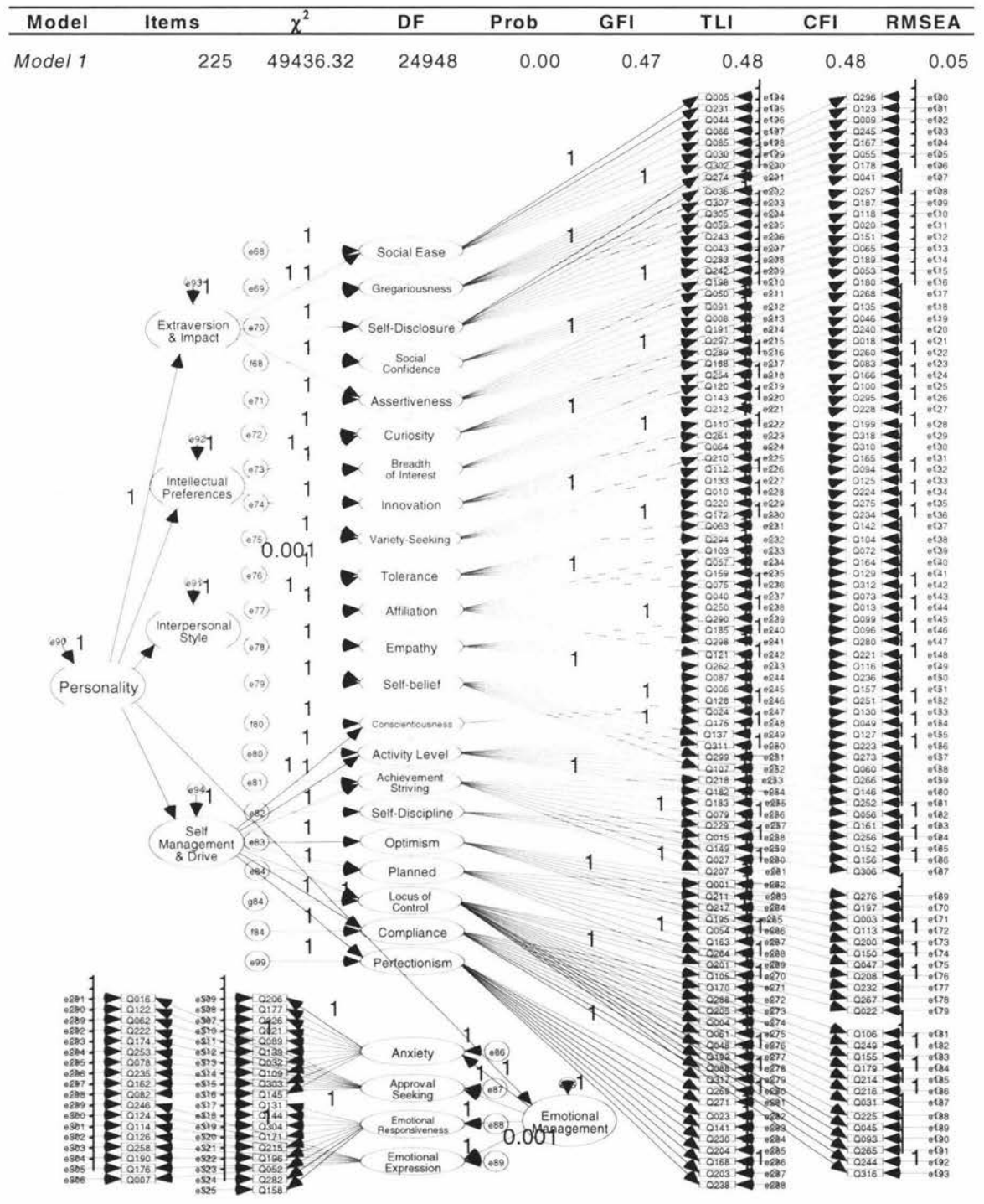
Fig. 9 showed the complete CFA for *Model 1* which is a fairly large model (225 items) for a dataset of only 484 participants. This means there were only two full participants per item, which is low, however the RMSEA was at a good level and therefore the analysis of the individual first-order factors could continue. As shown previously in Table 2, the fit statistics of many of the individual first-order factors used in this CFA did not reach acceptable levels of fit. Also shown in this table were the two first-order factors that did not converge.

The first, *Conscientiousness*, only had two items given that many of its original items cross-loaded and were therefore discarded. This meant that there were not enough parameters for this first-order CFA to converge. The second, *Achievement-Striving*, had 9 items however possibly due to the wording of these items this first-order factor also failed to converge. As Tabchnick and Fidell (2007) explain, for the benefit of fit and for explaining more of the variance associated with the latent variable, in this case *Personality*, it is preferred to retain existing factors in a model. This can be a trade-off and should not be performed at the cost of parsimony. In this case it was decided to re-include the cross-loaded items in order to assist with the convergence of these two first-order factors for the overall benefit of this questionnaire.

These two first-order factors had one item in common (Q315) meaning that it cross-loaded and therefore could not be included in the both CFAs. According to the *Squared Multiple Correlation* (SMC) this item was better suited to the *Achievement-Striving* first-order factor. These first-order factors were then reanalysed and successfully converged.

All 26 first-order factors were revised on the basis of the SMC. In most cases items were deleted if they were lower than .30 or when further deletions were necessary for the convergence of the CFA for each first-order factor. In regards to item Q315 referred to above, the SMC for *Conscientiousness* was below .30, and for *Achievement-Striving* was well above .30. Including the deletion of this item from the *Conscientiousness* first-order factor, Table 4 shows the 47 items that were deleted in order to satisfy the fit requirements referred to above.

Table 4

Items Deleted from Model 1 for Unidimensionality Analysis

Second-order Factor	First-order Factor	Deletions	Item	Squared Multiple Correlation
Extraversion and Impact				
	Gregariousness	3	Q009	0.27
			Q167	0.26
			Q245	0.11
	Self-Disclosure	2	Q242	0.13
			Q243	0.14
	Assertiveness	2	Q188	0.18
			Q289	0.21
Intellectual Preferences				
	Curiosity	1	Q135	0.21
	Breadth of Interest	3	Q143	0.13
			Q212	0.11
			Q254	0.08
	Variety-Seeking	2	Q010	0.08
			Q172	0.18
Interpersonal Style				
	Tolerance	1	Q275	0.06
	Affiliation	4	Q057	0.01
			Q103	0.02
			Q159	0.01
			Q250	0.05
	Empathy	1	Q013	0.10
	Self-belief	4	Q087	0.12
			Q128	0.12
			Q299	0.12
			Q311	0.10
Self Management and Drive				
	Conscientiousness	1	Q315	0.24
	Activity Level	3	Q252	0.11
			Q266	0.10
			Q273	0.06
	Self-Discipline	1	Q161	0.30
	Optimism	2	Q047	0.06
			Q077	0.11
	Planned	1	Q211	0.06
	Locus of Control	3	Q045	0.02
			Q214	0.07
			Q216	0.02
	Compliance	5	Q004	0.09
			Q048	0.19
			Q170	0.05
			Q193	0.12
			Q205	0.07
	Perfectionism	3	Q168	0.12
			Q203	0.15
			Q204	0.12
Emotional Management				
	Approval Seeking	3	Q082	0.00
			Q162	0.01
			Q174	0.12
	Emotional Expression	2	Q124	0.05
			Q126	0.01
Total Deletions		47		

The removal of poor items significantly improved the fit statistics for each of the individual first-order factors. The majority showed acceptable levels with only some that exceeded the RMSEA limit (see Table 5). This was the best output achievable as it was important to not delete too many of the items at this stage. These first-order factors ranged from 4 to 14 items each and the fit statistics and *Cronbach's Alpha* for each of the 24 first-order factors is shown in Table 5. These ranged from 0.63 to 0.86 ($M = 0.75$, $SD = .07$).

Table 5

Reliability and Fit Statistics for the 26 First-order Factors Ready for Model 2

Second-order Factor	First-order Factor	Items	Alpha	χ^2	DF	Prob	GFI	TLI	CFI	RMSEA
Extraversion and Impact										
	Social Ease	7	0.80	35.60	14	0.00	0.98	0.96	0.97	0.06
	Gregariousness	7	0.84	167.89	14	0.00	0.90	0.81	0.87	0.15
	Self-Disclosure	7	0.83	87.74	14	0.00	0.95	0.90	0.94	0.10
	Social-Confidence	6	0.75	132.44	9	0.00	0.92	0.69	0.81	0.17
	Assertiveness	6	0.79	72.69	9	0.00	0.95	0.86	0.92	0.12
Intellectual Preferences										
	Curiosity	7	0.82	73.75	14	0.00	0.96	0.91	0.94	0.09
	Breadth of Interest	4	0.69	17.63	2	0.00	0.98	0.89	0.96	0.13
	Innovation	9	0.80	192.67	27	0.00	0.91	0.80	0.85	0.11
	Variety-Seeking	6	0.77	44.93	9	0.00	0.97	0.91	0.95	0.09
Interpersonal Style										
	Tolerance	10	0.70	92.10	35	0.00	0.96	0.87	0.90	0.06
	Affiliation	6	0.66	35.92	9	0.00	0.98	0.87	0.92	0.08
	Empathy	6	0.76	29.93	9	0.00	0.97	0.93	0.95	0.07
	Self-belief	6	0.75	143.54	9	0.00	0.91	0.72	0.83	0.18
Self Management and Drive										
	Conscientiousness	9	0.79	95.21	27	0.00	0.96	0.90	0.93	0.07
	Activity Level	7	0.68	121.69	14	0.00	0.94	0.70	0.80	0.13
	Achievement-Striving	10	0.76	58.74	35	0.01	0.98	0.96	0.97	0.04
	Self-Discipline	4	0.76	11.24	2	0.00	0.99	0.94	0.98	0.10
	Optimism	6	0.65	112.99	9	0.00	0.92	0.65	0.79	0.16
	Planned	7	0.73	68.96	14	0.00	0.96	0.86	0.91	0.09
	Locus of Control	14	0.84	235.35	77	0.00	0.93	0.87	0.89	0.07
	Compliance	5	0.63	26.91	5	0.00	0.98	0.86	0.93	0.10
	Perfectionism	6	0.74	41.59	9	0.00	0.97	0.91	0.95	0.09
Emotional Management										
	Anxiety	10	0.83	144.72	35	0.00	0.94	0.89	0.92	0.08
	Approval Seeking	7	0.68	36.13	14	0.00	0.98	0.92	0.95	0.06
	Emotional Responsiveness	9	0.86	142.68	27	0.00	0.94	0.90	0.93	0.09
	Emotional Expression	6	0.64	24.21	9	0.00	0.98	0.93	0.96	0.06
Total Items		187								

Second-order CFA for Model 2

A CFA was then prepared for each second-order factor with their associated unidimensional first-order factors. The fit statistics for each of the five second-order factors is shown in Table 6 and the CFAs are shown in Appendix 4.

Table 6

Fit Statistics for the CFA of each Second-order Factor ready for Model 2

Second-order Factor	Items	χ^2	DF	Prob	GFI	TLI	CFI	RMSEA
Extraversion and Impact								
	33	1716.52	491	0.00	0.81	0.78	0.80	0.07
Intellectual Preferences								
	26	1144.79	295	0.00	0.83	0.78	0.80	0.08
Interpersonal Style								
	28	1032.88	347	0.00	0.86	0.77	0.79	0.06
Self Management and Drive								
	68	5292.42	2201	0.00	0.72	0.70	0.71	0.05
Emotional Management								
	32	1335.18	461	0.00	0.84	0.79	0.81	0.06
Total	**187**							

These second-order factors were then combined with the third-order factor *Personality* to form *Model 2*. The fit statistics for *Model 2* are shown in Fig. 10.

Fig. 10

Fit statistics for Model2

Model	Items	χ^2	DF	Prob	GFI	TLI	CFI	RMSEA
Model 2	187	33632.72	17174	0.00	0.53	0.57	0.58	0.05
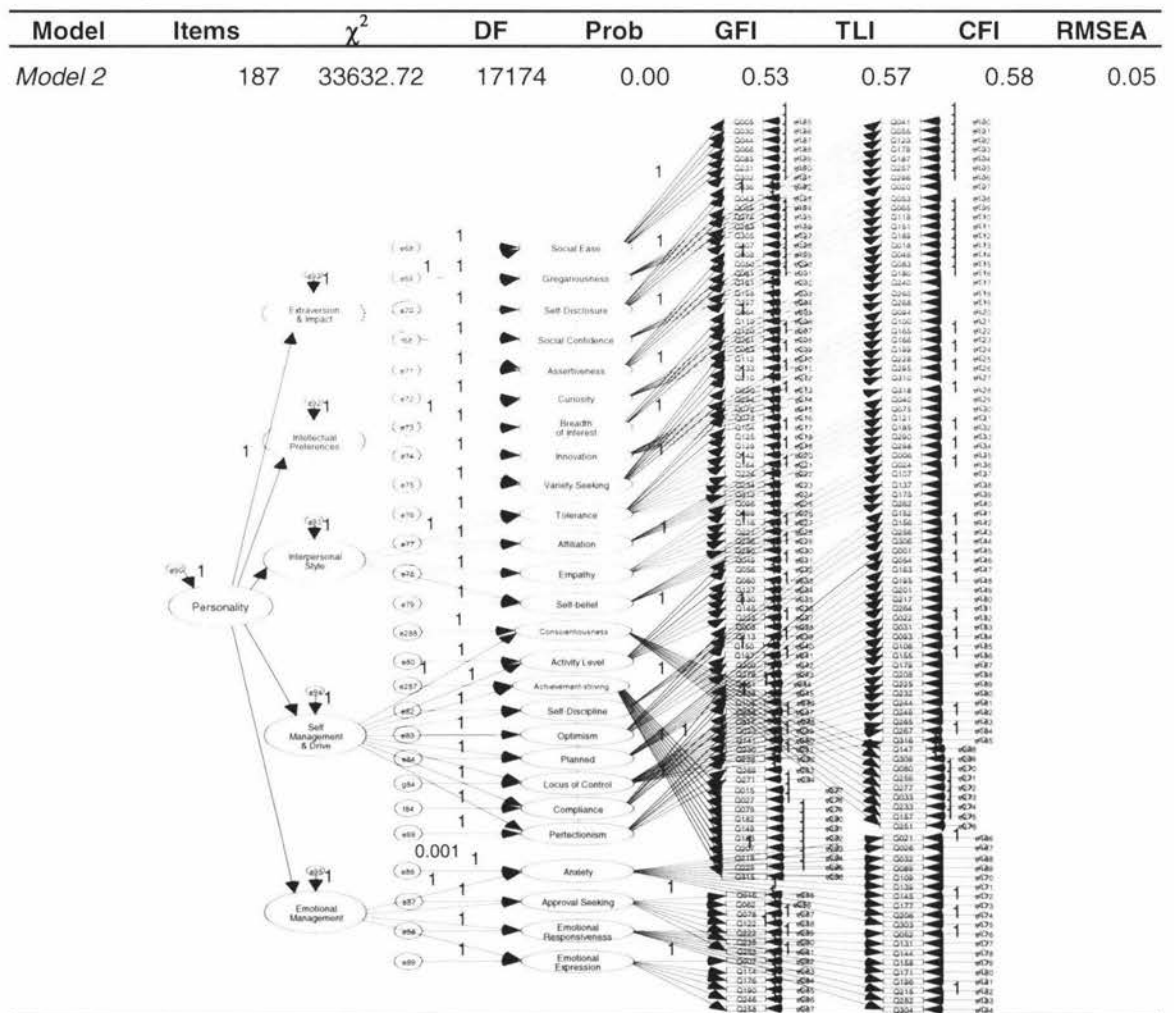
IRT Results

On the basis of the CFA results 187 items were available for IRT analysis. The '*a*' and '*b*' parameters were calculated for each of these items as shown in Table 7.

Table 7

Parameter Estimates for 187 Items in 26 First-order Factors

Second-order Factor	First-order Factor	Item	a	b1	b2	b3	b4	Wording
Extraversion and Impact								
	Social Ease							
		Q005	2.14	-3.35	-1.97	-0.96	0.72	Make friends easily.
		Q030	1.29	-5.13	-3.90	-2.66	-0.48	Am not really interested in others.
		Q044	1.29	-5.47	-3.17	-1.39	1.16	Cheer people up.
		Q066	1.99	-2.71	-1.11	-0.51	0.92	Am hard to get to know.
		Q085	1.41	-4.60	-2.71	-1.84	-0.15	Avoid contacts with others.
		Q231	1.94	-3.91	-1.96	-0.93	1.19	Warm up quickly to others.
		Q302	1.85	-3.73	-1.65	-0.83	0.82	Keep others at a distance.
	Gregariousness							
		Q041	2.16	-1.90	-0.61	0.12	1.12	Don't like crowded events.
		Q055	1.39	-3.99	-1.88	-0.25	1.39	Prefer to be alone.
		Q123	1.31	-2.77	-1.10	-0.34	1.53	Talk to a lot of different people at parties.
		Q178	1.32	-4.94	-2.40	-0.77	1.07	Want to be left alone.
		Q187	1.31	-3.35	-0.85	0.36	2.11	Seek quiet.
		Q257	2.62	-2.36	-1.00	-0.32	0.92	Avoid crowds.
		Q296	2.11	-1.77	-0.65	0.34	1.69	Love large parties.
	Self-Disclosure							
		Q036	2.05	-2.64	-1.39	-0.76	0.84	Am open about myself to others.
		Q043	2.00	-2.51	-1.00	-0.39	1.18	Reveal little about myself.
		Q059	1.45	-2.21	-0.66	0.12	2.14	Talk about my worries.
		Q274	3.46	-2.32	-1.00	-0.46	1.11	Am open about my feelings.
		Q283	1.03	-4.04	-1.57	-0.54	1.58	Bottle up my feelings.
		Q305	1.89	-2.62	-0.68	0.10	2.22	Show my feelings.
		Q307	1.42	-0.97	0.53	1.31	3.20	Disclose my intimate thoughts.
	Social-Confidence							
		Q020	1.41	-2.46	-1.18	-0.40	1.33	Don't mind being the centre of attention.
		Q053	1.83	-3.52	-2.58	-1.83	0.00	Am good at making impromptu speeches.
		Q065	2.00	-3.24	-1.96	-0.91	0.99	Express myself easily.
		Q118	1.35	-2.61	-1.17	-0.30	1.82	Have a natural talent for influencing people.
		Q151	1.57	-4.11	-2.46	-1.64	0.56	Hate being the centre of attention.
		Q189	1.24	-2.64	-0.95	0.10	2.24	Lack the talent for influencing people.
	Assertiveness							
		Q008	1.15	-4.74	-3.35	-1.91	0.62	Seek to influence others.
		Q050	2.08	-3.45	-2.70	-1.60	0.18	Try to lead others.
		Q091	1.37	-4.54	-3.18	-1.81	1.06	Can talk others into doing things.
		Q191	2.16	-4.19	-2.58	-1.62	0.62	Take control of things.
		Q198	3.24	-3.59	-2.17	-1.33	0.56	Take charge.
		Q297	1.53	-4.35	-2.29	-1.23	0.89	Wait for others to lead the way.
Intellectual Preferences								
	Curiosity							
		Q018	1.77	-3.07	-1.75	-0.86	0.55	Avoid philosophical discussions.
		Q046	1.38	-4.34	-3.76	-2.64	-0.45	Enjoy thinking about things.
		Q083	2.24	-2.91	-1.71	-1.08	0.43	Am not interested in theoretical discussions.
		Q180	1.11	-5.49	-3.79	-2.46	-0.01	Like to solve complex problems.
		Q240	2.66	-2.71	-1.79	-1.17	0.23	Am not interested in abstract ideas.
		Q260	2.08	-3.07	-1.92	-1.21	0.29	Have difficulty understanding abstract ideas.
		Q268	1.06	-4.95	-2.60	-1.02	1.32	Have a rich vocabulary.

Table 7 *(continued)*

Second-order Factor	First-order Factor	Item	a	b1	b2	b3	b4	Wording
	Breadth of Interest							
		Q064	0.79	-5.56	-3.54	-1.89	0.24	Do not like concerts.
		Q110	2.13	-2.23	-1.35	-0.76	0.33	Do not enjoy going to art museums.
		Q120	0.74	-8.78	-5.76	-4.32	-0.51	Am interested in many things.
		Q261	5.79	-2.55	-1.68	-1.07	-0.15	Do not like art.
	Innovation							
		Q094	2.09	-3.69	-2.41	-2.04	-0.50	Can't come up with new ideas.
		Q100	2.27	-3.39	-2.23	-1.21	0.46	Am full of ideas.
		Q165	0.92	-11.51	-3.57	-2.41	0.58	Will not probe deeply into a subject.
		Q166	1.35	-3.36	-1.99	-0.69	1.12	Have a vivid imagination.
		Q199	1.92	-4.45	-2.90	-1.32	0.75	Have excellent ideas.
		Q228	1.36	-3.65	-1.92	-0.60	1.74	Come up with bold plans.
		Q295	1.00	-5.38	-3.04	-0.45	2.50	Carry the conversation to a higher level.
		Q310	1.92	-3.82	-2.56	-1.82	-0.06	Have difficulty imagining things.
		Q318	1.81	-3.16	-2.27	-1.71	-0.27	Do not have a good imagination.
	Variety-Seeking							
		Q063	2.55	-3.01	-2.09	-1.36	0.05	Don't like the idea of change.
		Q112	2.18	-2.83	-2.56	-2.11	-0.12	Am open to change.
		Q133	1.50	-7.61	-4.30	-3.26	-0.89	Enjoy hearing new ideas.
		Q210	1.21	-5.08	-2.87	-1.61	0.36	Prefer variety to routine.
		Q220	1.48	-4.84	-3.31	-1.96	0.53	Like to begin new things.
		Q294	2.80	-2.86	-2.09	-1.41	0.22	Dislike changes.
Interpersonal Style								
	Tolerance							
		Q072	1.19	-4.68	-1.83	-0.93	0.92	Get irritated easily.
		Q073	0.98	-3.84	-0.46	0.78	3.17	Am annoyed by others' mistakes.
		Q104	0.75	-4.29	-1.88	-0.71	1.64	Am a bad loser.
		Q125	1.23	-6.19	-2.74	-1.82	0.63	Accept people as they are.
		Q129	1.95	-3.91	-2.80	-2.02	-0.09	Treat people as inferiors.
		Q142	0.80	-12.39	-5.41	-2.60	1.88	Believe that others have good intentions.
		Q164	0.51	-8.14	-2.55	0.16	3.66	Lay down the law to others.
		Q224	1.48	-4.91	-4.15	-2.87	-0.39	Respect others.
		Q234	0.68	-8.43	-5.94	-4.28	-0.30	Believe there are many sides to most issues.
		Q312	1.27	-3.13	-1.15	-0.28	1.74	Am quick to judge others.
	Affiliation							
		Q040	0.94	-4.78	-1.61	0.23	2.30	Contradict others.
		Q075	1.68	-2.83	-1.14	-0.44	0.57	Have a sharp tongue.
		Q121	1.02	-5.46	-2.60	-1.56	0.26	Hold a grudge.
		Q185	2.26	-3.45	-2.60	-2.02	-1.06	Insult people.
		Q290	1.33	-5.29	-3.37	-2.25	-0.88	Yell at people.
		Q298	1.00	-1.39	-0.41	0.41	1.39	Get back at others.
	Empathy							
		Q096	1.68	-4.23	-2.96	-1.32	1.44	Make others feel good.
		Q099	1.94	-4.18	-3.67	-2.14	0.27	Reassure others.
		Q116	2.14	-6.26	-3.21	-1.87	0.21	Make people feel welcome.
		Q221	0.99	-5.07	-2.54	-0.79	2.37	Have a good word for everyone.
		Q236	2.23	-6.17	-2.54	-1.49	0.48	Take time out for others.
		Q280	1.71	-7.23	-3.39	-2.12	0.50	Am concerned about others.
	Self-belief							
		Q006	0.86	-2.84	-0.40	0.52	2.54	Consider myself an average person.
		Q024	1.07	-4.09	-1.10	0.67	1.87	Believe that I am better than others.
		Q107	0.73	-5.91	-2.47	-0.57	1.86	Make myself the center of attention.
		Q137	4.41	-1.39	0.06	0.84	1.77	Have a high opinion of myself.
		Q175	3.18	-1.03	0.47	1.32	2.12	Think highly of myself.
		Q262	0.68	-3.85	-0.31	1.24	4.39	Dislike being the center of attention.

Table 7 *(continued)*

Second-order Factor	First-order Factor	Item	a	b1	b2	b3	b4	Wording
Self Management and Drive								
	Conscientiousness							
		Q033	1.61	-3.49	-2.09	-1.16	0.25	Waste my time.
		Q080	1.25	-3.81	-1.38	-0.56	1.76	Get chores done right away.
		Q147	1.35	-4.59	-2.10	-1.04	1.89	Am always prepared.
		Q157	1.73	-3.41	-2.09	-1.64	-0.17	Don't see things through.
		Q233	2.00	-3.33	-2.31	-1.53	0.02	Find it difficult to get down to work.
		Q251	1.89	-3.45	-3.07	-2.58	-1.31	Shirk my duties.
		Q259	2.27	-4.02	-3.40	-2.34	0.16	Carry out my plans.
		Q277	1.66	-4.49	-2.53	-1.51	1.40	Make plans and stick to them.
		Q309	1.19	-4.82	-2.49	-1.63	0.88	Pay attention to details.
	Activity Level							
		Q049	2.17	-3.15	-1.77	-1.05	0.62	Am always on the go.
		Q056	0.49	-9.80	-4.51	-2.19	1.68	React slowly.
		Q060	0.92	-4.41	-1.40	-0.23	1.89	Like to take it easy.
		Q127	1.68	-3.55	-2.07	-1.45	0.24	Do a lot in my spare time.
		Q130	2.46	-3.12	-1.86	-1.23	0.41	Am always busy.
		Q146	0.86	-3.48	-0.83	0.30	3.07	Like a leisurely lifestyle.
		Q223	0.76	-8.49	-4.96	-3.36	0.41	Can manage many things at the same time.
	Achievement-Striving							
		Q015	1.80	-3.99	-3.46	-2.65	-0.54	Set high standards for myself and others.
		Q027	1.25	-3.48	-2.82	-2.24	-0.66	Am not highly motivated to succeed.
		Q079	1.07	-5.20	-3.12	-1.69	0.95	Plunge into tasks with all my heart.
		Q149	1.61	-4.65	-3.63	-2.30	0.11	Demand quality.
		Q182	2.25	-3.69	-3.23	-2.55	-0.75	Work hard.
		Q183	2.12	-6.24	-3.11	-2.32	-0.36	Turn plans into actions.
		Q207	1.88	-3.11	-2.63	-2.45	-1.32	Put little time and effort into my work.
		Q218	0.90	-6.92	-3.13	-1.29	2.01	Go straight for the goal.
		Q229	1.55	-4.80	-3.54	-1.88	0.38	Do more than what's expected of me.
		Q315	1.90	-4.02	-2.94	-2.30	0.60	Do just enough work to get by.
	Self-Discipline							
		Q152	2.86	-2.97	-1.68	-1.11	0.10	Need a push to get started.
		Q156	2.89	-2.75	-1.60	-1.13	0.22	Have difficulty starting tasks.
		Q256	1.39	-4.25	-2.25	-1.20	1.05	Get to work at once.
		Q306	1.60	-3.64	-1.92	-0.99	0.99	Postpone decisions.
	Optimism							
		Q003	1.20	-5.40	-3.85	-2.39	0.56	Know what I want.
		Q113	1.08	-3.91	-1.91	-1.23	0.76	Am not sure where my life is going.
		Q150	1.02	-6.55	-4.36	-2.62	0.00	Let others determine my choices.
		Q197	2.69	-3.27	-2.46	-1.66	0.21	Keep improving myself.
		Q200	0.92	-5.67	-3.23	-2.15	-0.53	Am resigned to my fate.
		Q276	2.95	-5.40	-2.71	-1.94	-0.13	Work on improving myself.
	Planned							
		Q001	0.94	-7.04	-4.73	-3.46	0.31	Like to plan ahead.
		Q054	1.26	-3.78	-1.47	0.33	1.27	Like to act on a whim.
		Q163	1.06	-2.89	-0.92	0.18	2.63	Often make last-minute plans.
		Q195	2.53	-3.35	-1.71	-1.23	0.31	Jump into things without thinking.
		Q201	1.43	-5.53	-3.96	-2.53	-0.83	Make a mess of things.
		Q217	0.85	-6.21	-3.38	-1.56	1.70	Am exacting in my work.
		Q264	2.14	-3.76	-2.30	-1.61	0.09	Make rash decisions.

Table 7 *(continued)*

Second-order Factor	First-order Factor	Item	a	b1	b2	b3	b4	Wording
	Locus of Control							
		Q022	1.38	-5.17	-4.16	-2.83	0.32	Come up with good solutions.
		Q031	1.05	-5.77	-3.25	-1.70	0.24	See difficulties everywhere.
		Q093	1.92	-3.17	-2.50	-2.15	-0.76	Dislike taking responsibility for making decisions.
		Q106	1.33	-4.48	-2.86	-1.92	0.75	Feel up to any task.
		Q155	1.86	-4.62	-3.52	-2.21	-0.13	Take the initiative.
		Q179	1.29	-5.97	-2.93	-2.09	0.72	Make a decision and move on.
		Q208	1.19	-4.29	-3.36	-2.69	-0.10	Feel comfortable with myself.
		Q225	1.38	-5.67	-3.42	-2.38	-0.67	Habitually blow my chances.
		Q232	1.11	-5.78	-4.50	-3.12	-0.23	Believe that my success depends on ability rather than luck.
		Q244	1.55	-4.32	-3.35	-2.44	-0.97	Dislike myself.
		Q249	1.76	-3.98	-3.30	-2.42	-0.36	Like to take responsibility for making decisions.
		Q265	1.35	-4.77	-4.07	-2.63	-0.64	Am less capable than most people.
		Q267	0.97	-4.36	-2.56	-0.73	1.73	Just know that I will be a success.
		Q316	1.86	-3.48	-2.80	-2.35	-0.67	Feel that I'm unable to deal with things.
	Compliance							
		Q061	1.77	-3.09	-1.00	-0.23	1.14	Break rules.
		Q088	1.33	-4.96	-4.04	-2.49	-0.87	Do the opposite of what is asked.
		Q105	2.47	-3.17	-1.90	-1.22	0.50	Try to follow the rules.
		Q288	0.71	-7.51	-4.23	-3.09	-0.22	Pay my bills on time.
		Q317	0.88	-6.18	-5.05	-4.04	-1.63	Misrepresent the facts.
	Perfectionism							
		Q023	1.79	-3.09	-1.51	-0.62	1.27	Want everything to be "just right."
		Q141	1.08	-4.84	-2.34	-1.19	2.04	Want things to proceed according to plan.
		Q230	1.26	-2.82	-1.07	0.09	2.84	Demand perfection in others.
		Q238	0.77	-3.81	-1.14	-0.16	2.59	Am not bothered by disorder.
		Q269	2.40	-2.05	-0.60	0.11	2.03	Continue until everything is perfect.
		Q271	2.80	-1.94	-0.64	-0.14	1.49	Want every detail taken care of.
Emotional Management								
	Anxiety							
		Q021	1.66	0.00	1.25	2.19	3.78	Have frequent mood swings.
		Q026	2.01	-0.59	0.97	1.81	3.51	Get upset easily.
		Q032	1.33	-1.68	0.48	1.38	3.66	Am not easily bothered by things.
		Q089	1.44	0.20	1.52	2.74	4.54	Often feel blue.
		Q109	1.62	-1.52	0.36	0.95	2.93	Rarely get irritated.
		Q139	1.28	-1.50	1.08	1.89	4.05	Am relaxed most of the time.
		Q145	1.61	-1.47	0.59	1.23	3.25	Am not easily frustrated.
		Q177	1.14	-2.84	-0.68	0.22	2.89	Worry about things.
		Q206	2.11	-0.43	1.10	1.80	3.14	Get stressed out easily.
		Q303	1.14	-0.83	0.97	1.64	3.35	Seldom feel blue.
	Approval Seeking							
		Q016	1.97	-2.46	-1.09	-0.28	1.64	Worry about what people think of me.
		Q062	1.86	-1.81	-0.45	0.34	2.43	Need the approval of others.
		Q078	0.99	-4.18	-1.97	-1.21	1.44	Am not concerned with making a good impression.
		Q122	0.74	-2.03	1.12	3.47	8.18	Conform to others' opinions.
		Q222	1.13	-3.52	-1.85	-0.55	1.61	Want to amount to something special in others' eyes.
		Q235	0.76	-0.80	2.26	3.22	6.51	Feel it's OK that some people don't like me.
		Q253	1.10	-4.67	-1.67	-0.72	1.81	Don't care what others think.

Table 7 *(continued)*

Second-order Factor	First-order Factor	Item	a	b1	b2	b3	b4	Wording
	Emotional Responsiveness							
		Q052	1.76	-2.81	-1.34	-0.70	0.99	Rarely analyse my emotions.
		Q131	2.30	-2.69	-1.42	-0.85	0.84	Think about the causes of my emotions.
		Q144	2.17	-2.79	-1.12	-0.07	1.62	Pay a lot of attention to my feelings.
		Q158	1.45	-3.97	-1.59	-0.58	1.41	Often ignore my feelings.
		Q171	3.24	-3.19	-1.69	-0.95	0.79	Notice my emotions.
		Q196	2.26	-3.00	-1.54	-0.85	0.88	Rarely think about how I feel.
		Q215	1.59	-2.21	-0.48	0.44	2.37	Often stop to analyse how I'm feeling.
		Q282	1.38	-3.80	-2.32	-1.45	0.64	Am not in touch with my feelings.
		Q304	1.27	-4.65	-3.46	-2.19	0.72	Am usually aware of the way that I'm feeling.
	Emotional Expression							
		Q007	0.68	-6.90	-3.23	-1.80	1.63	Don't understand people who get emotional.
		Q114	0.29	-13.43	-6.79	-2.88	3.64	Enjoy examining myself and my life.
		Q176	1.23	-2.12	-0.21	0.58	2.84	Experience very few emotional highs and lows.
		Q190	1.91	-1.86	0.03	0.74	2.68	Am not easily affected by my emotions.
		Q246	1.35	-1.59	0.02	1.09	3.03	Experience my emotions intensely.
		Q258	1.93	-1.37	0.28	0.89	2.54	Seldom get emotional.

Fig. 11 shows the *Test Information Function* (TIF) for *Model 2*. The information captured by this model is relatively high however it does operate largely in the lower end of the personality continuum.

Fig. 11

Test Information Function for Model 2 (187 items)
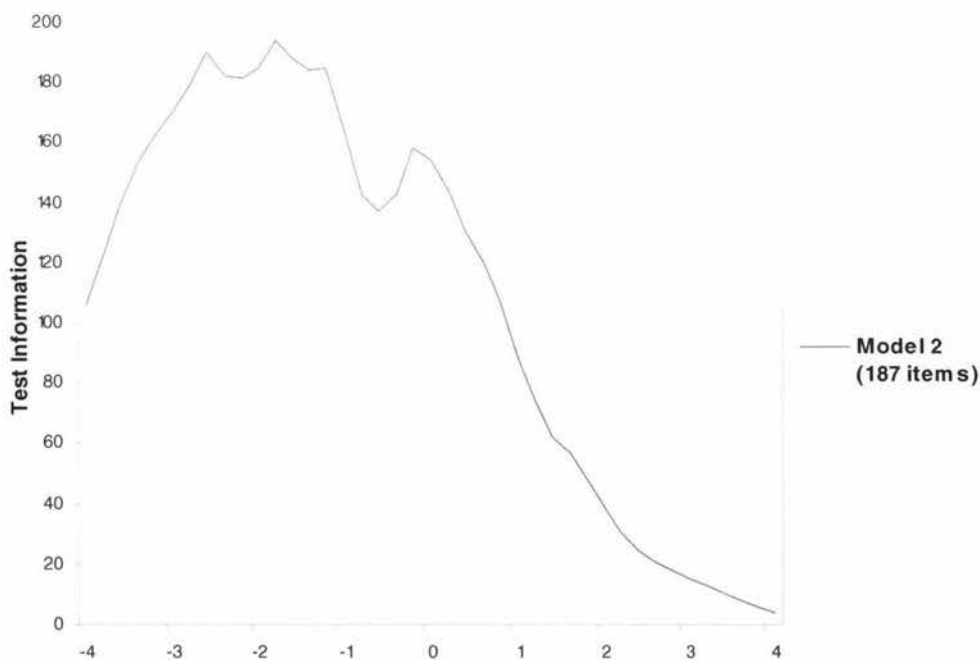
Scale Information Functions were calculated to show which factors had the most information and where it was located (see Fig. 12).

Fig. 12

Scale Information Function for all Five Second-order Factors for Model 2
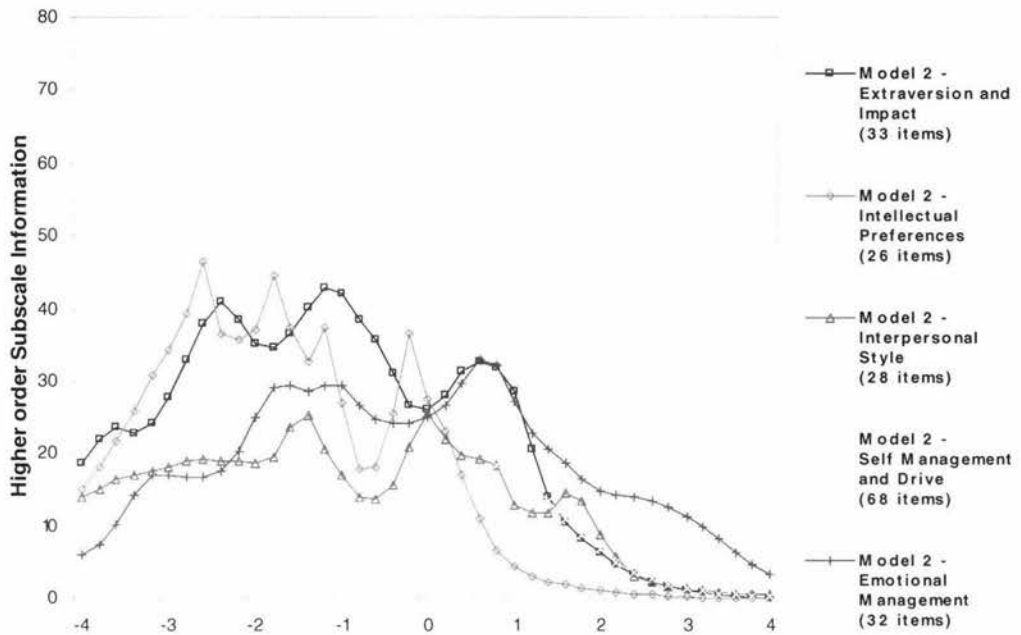
Item Selection Summary

Of the 187 items analysed in this section and using the criteria stated earlier, three items were identified for all 26 first-order factors, meaning 69 were selected and 118 were discarded. Therefore, the objective of refining each first-order factor from its original length down to three items was satisfied.

To give an indication of the overall item quality in this questionnaire, the *Model 2* items were categorised into 'poor', 'mediocre', and 'good' item categories using the IRT properties mentioned previously. There were a total of 187 individual items and these were categorised as follows: *poor items*- 65, *mediocre items*- 31, and *good items*- 91. The vast majority of selections for *Model 3* came from the *good items* category with very few added from the *mediocre items*. These *mediocre* items were used at times when item combinations were valued over higher '*a*' parameters, in order to maximise the variance of the item wording in each first-order factor.

Along with the *Cronbach's Alpha* for each first-order factor, the items in the final selection and their '*a*' and '*b*' parameters are shown in Table 8. The corresponding CCC's, OCF's, and IIF's for the 78 items selected for *Model 3* are shown in Appendix 5.

Table 8

Parameter Estimates for the 78 Items in Model 3 and Cronbach's Alpha each First-order Factor

Second-order Factor	First-order Factor	Alpha	Item	a	b1	b2	b3	b4	Wording
Extraversion & Impact									
	Social Ease								
		0.76	Q005	2.14	-3.35	-1.97	-0.96	0.72	Make friends easily.
			Q066	1.99	-2.71	-1.11	-0.51	0.92	Am hard to get to know.
			Q231	1.94	-3.91	-1.96	-0.93	1.19	Warm up quickly to others.
	Gregariousness								
		0.82	Q041	2.16	-1.90	-0.61	0.12	1.12	Don't like crowded events.
			Q257	2.62	-2.36	-1.00	-0.32	-0.92	Avoid crowds.
			Q296	2.11	-1.77	-0.65	0.34	1.69	Love large parties.
	Self-Disclosure								
		0.77	Q036	2.05	-2.64	-1.39	-0.76	0.84	Am open about myself to others.
			Q274	3.46	-2.32	-1.00	-0.46	1.11	Am open about my feelings.
			Q305	1.89	-2.62	-0.68	0.10	2.22	Show my feelings.
	Social-Confidence								
		0.67	Q020	1.41	-2.46	-1.18	-0.40	1.33	Don't mind being the centre of attention.
			Q053	1.83	-3.52	-2.58	-1.83	0.00	Am good at making impromptu speeches.
			Q065	2.00	-3.24	-1.96	-0.91	0.99	Express myself easily.
	Assertiveness								
		0.76	Q050	2.08	-3.45	-2.70	-1.60	0.18	Try to lead others.
			Q191	2.16	-4.19	-2.58	-1.62	0.62	Take control of things.
			Q198	3.24	-3.59	-2.17	-1.33	0.56	Take charge.
Intellectual Preferences									
	Curiosity								
		0.75	Q018	1.77	-3.07	-1.75	-0.86	0.55	Avoid philosophical discussions.
			Q083	2.24	-2.91	-1.71	-1.08	0.43	Am not interested in theoretical discussions.
			Q240	2.66	-2.71	-1.79	-1.17	0.23	Am not interested in abstract ideas.
	Breadth of Interest								
		0.65	Q064	0.79	-5.56	-3.54	-1.89	0.24	Do not like concerts.
			Q110	2.13	-2.23	-1.35	-0.76	0.33	Do not enjoy going to art museums.
			Q261	5.79	-2.55	-1.68	-1.07	-0.15	Do not like art.
	Innovation								
		0.71	Q100	2.27	-3.39	-2.23	-1.21	0.46	Am full of ideas.
			Q199	1.92	-4.45	-2.90	-1.32	0.75	Have excellent ideas.
			Q228	1.36	-3.65	-1.92	-0.60	1.74	Come up with bold plans.
	Variety-Seeking								
		0.67	Q063	2.55	-3.01	-2.09	-1.36	0.05	Don't like the idea of change.
			Q220	1.48	-4.84	-3.31	-1.96	0.53	Like to begin new things.
			Q294	2.80	-2.86	-2.09	-1.41	0.22	Dislike changes.
Interpersonal Style									
	Tolerance								
		0.55	Q125	1.23	-6.19	-2.74	-1.82	0.63	Accept people as they are.
			Q129	1.95	-3.91	-2.80	-2.02	-0.09	Treat people as inferiors.
			Q224	1.48	-4.91	-4.15	-2.87	-0.39	Respect others.
	Affiliation								
		0.53	Q075	1.68	-2.83	-1.14	-0.44	0.57	Have a sharp tongue.
			Q185	2.26	-3.45	-2.60	-2.02	-1.06	Insult people.
			Q290	1.33	-5.29	-3.37	-2.25	-0.88	Yell at people.
	Empathy								
		0.71	Q116	2.14	-6.26	-3.21	-1.87	0.21	Make people feel welcome.
			Q236	2.23	-6.17	-2.54	-1.49	0.48	Take time out for others.
			Q280	1.71	-7.23	-3.39	-2.12	0.50	Am concerned about others.
	Self-belief								
		0.73	Q024	1.07	-4.09	-1.10	0.67	1.87	Believe that I am better than others.
			Q137	4.41	-1.39	0.06	0.84	1.77	Have a high opinion of myself.
			Q175	3.18	-1.03	0.47	1.32	2.12	Think highly of myself.

Table 8 *(continued)*

Second-order Factor	First-order Factor	Alpha	Item	a	b1	b2	b3	b4	Wording
Self Management & Drive									
	Conscientiousness								
		0.67	Q033	1.61	-3.49	-2.09	-1.16	0.25	Waste my time.
			Q233	2.00	-3.33	-2.31	-1.53	0.02	Find it difficult to get down to work.
			Q259	2.27	-4.02	-3.40	-2.34	0.16	Carry out my plans.
	Activity Level								
		0.73	Q049	2.17	-3.15	-1.77	-1.05	0.62	Am always on the go.
			Q127	1.68	-3.55	-2.07	-1.45	0.24	Do a lot in my spare time.
			Q130	2.46	-3.12	-1.86	-1.23	0.41	Am always busy.
	Achievement-Striving								
		0.61	Q149	1.61	-4.65	-3.63	-2.30	0.11	Demand quality.
			Q229	1.55	-4.80	-3.54	-1.88	0.38	Do more than what's expected of me.
			Q315	1.90	-4.02	-2.94	-2.30	0.60	Do just enough work to get by.
	Self-Discipline								
		0.68	Q152	2.86	-2.97	-1.68	-1.11	0.10	Need a push to get started.
			Q256	1.39	-4.25	-2.25	-1.20	1.05	Get to work at once.
			Q306	1.60	-3.64	-1.92	-0.99	0.99	Postpone decisions.
	Optimism								
		0.71	Q003	1.20	-5.40	-3.85	-2.39	0.56	Know what I want.
			Q197	2.69	-3.27	-2.46	-1.66	0.21	Keep improving myself.
			Q276	2.95	-5.40	-2.71	-1.94	-0.13	Work on improving myself.
	Planned								
		0.68	Q054	1.26	-3.78	-1.47	0.33	1.27	Like to act on a whim.
			Q195	2.53	-3.35	-1.71	-1.23	0.31	Jump into things without thinking.
			Q264	2.14	-3.76	-2.30	-1.61	0.09	Make rash decisions.
	Locus of Control								
		0.58	Q225	1.44	-5.52	-3.36	-2.36	-0.69	Habitually blow my chances.
			Q244	1.68	-4.14	-3.24	-2.38	-0.97	Dislike myself.
			Q249	1.65	-4.20	-3.47	-2.55	-0.41	Like to take responsibility for making decisions.
	Compliance								
		0.60	Q061	1.77	-3.09	-1.00	-0.23	1.14	Break rules.
			Q088	1.33	-4.96	-4.04	-2.49	-0.87	Do the opposite of what is asked.
			Q105	2.47	-3.17	-1.90	-1.22	0.50	Try to follow the rules.
	Perfectionism								
		0.77	Q023	1.79	-3.09	-1.51	-0.62	1.27	Want everything to be "just right."
			Q269	2.40	-2.05	-0.60	0.11	2.03	Continue until everything is perfect.
			Q271	2.80	-1.94	-0.64	-0.14	1.49	Want every detail taken care of.
Emotional Management									
	Anxiety								
		0.72	Q026	2.01	-0.59	0.97	1.81	3.51	Get upset easily.
			Q032	1.33	-1.68	0.48	1.38	3.66	Am not easily bothered by things.
			Q206	2.11	-0.43	1.10	1.80	3.14	Get stressed out easily.
	Approval Seeking								
		0.64	Q016	1.97	-2.46	-1.09	-0.28	1.64	Worry about what people think of me.
			Q062	1.86	-1.81	-0.45	0.34	2.43	Need the approval of others.
			Q222	1.13	-3.52	-1.85	-0.55	1.61	Want to amount to something special in others' eyes.
	Emotional Responsiveness								
		0.84	Q131	2.30	-2.69	-1.42	-0.85	0.84	Think about the causes of my emotions.
			Q144	2.17	-2.79	-1.12	-0.07	1.62	Pay a lot of attention to my feelings.
			Q171	3.24	-3.19	-1.69	-0.95	0.79	Notice my emotions.
	Emotional Expression								
		0.68	Q190	1.91	-1.86	0.03	0.74	2.68	Am not easily affected by my emotions.
			Q246	1.35	-1.59	0.02	1.09	3.03	Experience my emotions intensely.
			Q258	1.93	-1.37	0.28	0.89	2.54	Seldom get emotional.

These 78 items were then re-grouped by their associated second-order factor and the five CFAs were run with the second half of the dataset ($n = 489$). This was performed as part of the process of confirming the factor structure of this measure and also to ensure a valid test of this reduced length measure at the factor level before combining the factors into a 3-stage higher-order model.

Fig. 13 illustrates the item selection process by showing all three CFAs for one of the factors, *Interpersonal Style*. The first stage shows the second-order CFA from *Model 1*, the second stage from *Model 2*, and the third from *Model 3*. As can be seen in the table in Fig. 13, the deletion of poor items has improved the fit statistics considerably throughout the transition between these stages.

Fig. 13

Fit Statistics and CFAs for the Development Process for the Second-order Factor Interpersonal Style

Stage	Items	χ^2	DF	Prob	GFI	TLI	CFI	RMSEA
1	38	2085.83	662	0.00	0.79	0.64	0.66	0.07
2	28	1032.88	347	0.00	0.86	0.77	0.79	0.06
3	12	131.51	51	0.00	0.96	0.93	0.94	0.06
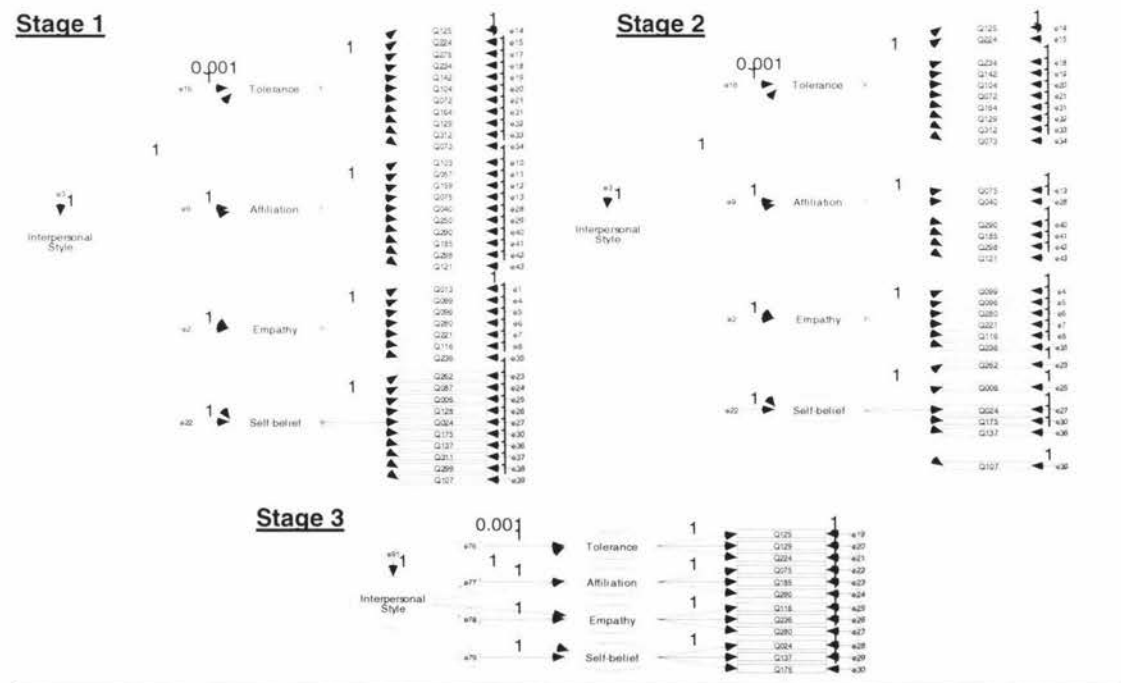
The fit statistics for the comparisons between the three models for each individual factor are shown in Table 9.

Table 9

Fit Statistics for the Development Process for all Five Second-order Factors

Second-order Factor	Model	Items	χ^2	DF	Prob	GFI	TLI	CFI	RMSEA
Extraversion and Impact									
	Model 1	40	2886.44	735	0.00	0.74	0.70	0.72	0.08
	Model 2	33	1716.52	491	0.00	0.81	0.78	0.80	0.07
	Model 3	15	286.05	85	0.00	0.93	0.90	0.92	0.07
Intellectual Preferences									
	Model 1	32	1748.09	460	0.00	0.79	0.73	0.75	0.08
	Model 2	26	1144.79	295	0.00	0.83	0.78	0.80	0.08
	Model 3	12	163.22	50	0.00	0.95	0.91	0.93	0.07
Interpersonal Style									
	Model 1	38	2085.83	662	0.00	0.79	0.64	0.66	0.07
	Model 2	28	1032.88	347	0.00	0.86	0.77	0.79	0.06
	Model 3	12	131.51	51	0.00	0.96	0.93	0.94	0.06
Self Management and Drive									
	Model 1	78	6689.77	2916	0.00	0.70	0.64	0.65	0.05
	Model 2	68	5292.42	2201	0.00	0.72	0.70	0.71	0.05
	Model 3	27	938.74	316	0.00	0.87	0.83	0.85	0.06
Emotional Management									
	Model 1	37	1905.73	626	0.00	0.79	0.73	0.75	0.07
	Model 2	32	1335.18	461	0.00	0.84	0.79	0.81	0.06
	Model 3	12	168.88	51	0.00	0.95	0.91	0.93	0.07

Reliability comparisons were made between the first-order factors for each of the three models. These are shown in Table 10 along with the average reliability for each model shown at the bottom of the table. The majority of the *Cronbach's Alpha*s for *Model 3* are reasonably high considering that there are only three items in each of the first-order factors. However, due to reasons explained earlier, obtaining high first-order factor reliability is not an aim of this selection process.

Table 10

Reliability Comparisons for the First-order Factors between Model 1, Model 2 and Model 3

Second-order Factor	First-order Factor	Model 1		Model 2		Model 3	
		Items	Alpha	Items	Alpha	Items	Alpha
Extraversion and Impact							
	Social Ease	7	0.80	7	0.80	3	0.76
	Gregariousness	10	0.85	7	0.84	3	0.82
	Self-Disclosure	9	0.83	7	0.83	3	0.77
	Social-Confidence	6	0.75	6	0.75	3	0.67
	Assertiveness	8	0.79	6	0.79	3	0.76
		40		33		15	
Intellectual Preferences							
	Curiosity	8	0.83	7	0.82	3	0.75
	Breadth of Interest	7	0.69	4	0.69	3	0.65
	Innovation	9	0.80	9	0.80	3	0.71
	Variety-Seeking	8	0.76	6	0.77	3	0.67
		32		26		12	
Interpersonal Style							
	Tolerance	11	0.69	10	0.70	3	0.55
	Affiliation	10	0.59	6	0.66	3	0.53
	Empathy	7	0.75	6	0.76	3	0.71
	Self-belief	10	0.78	6	0.75	3	0.73
		38		28		12	
Self Management and Drive							
	Conscientiousness	2	0.36	9	0.79	3	0.67
	Activity Level	10	0.71	7	0.68	3	0.73
	Achievement-Striving	9	0.73	10	0.76	3	0.61
	Self-Discipline	5	0.79	4	0.76	3	0.68
	Optimism	8	0.66	6	0.65	3	0.71
	Planned	8	0.72	7	0.73	3	0.68
	Locus of Control	17	0.80	14	0.84	3	0.58
	Compliance	10	0.68	5	0.63	3	0.60
	Perfectionism	9	0.76	6	0.74	3	0.77
		78		68		27	
Emotional Management							
	Anxiety	10	0.83	10	0.83	3	0.72
	Approval Seeking	10	0.65	7	0.68	3	0.64
	Emotional Responsiveness	9	0.86	9	0.86	3	0.84
	Emotional Expression	8	0.64	6	0.64	3	0.68
		37		32		12	
		Total	Average	Total	Average	Total	Average
		225	0.73	187	0.75	78	0.69

A premise of this study was to reconstitute this personality questionnaire by reducing it to only the items that capture high information, thus making the test more precise. The increased precision also meant a slight reduction of information. This was assessed at the second-order factor level by calculating a *Scale Information Function* (SIF) for *Model 2* and *Model 3* on the same graph. This is similar to the IIF's shown previously but illustrates the information captured by the whole factor rather than the individual item. The SIF for the second-order factor '*Intellectual Preferences*' is shown below in Fig. 14 and illustrates the change in information from *Model 2* to *Model 3*.

Fig. 14

Comparison of the SIF for the Second-order Factor Intellectual Preferences for Model 2 and Model 3
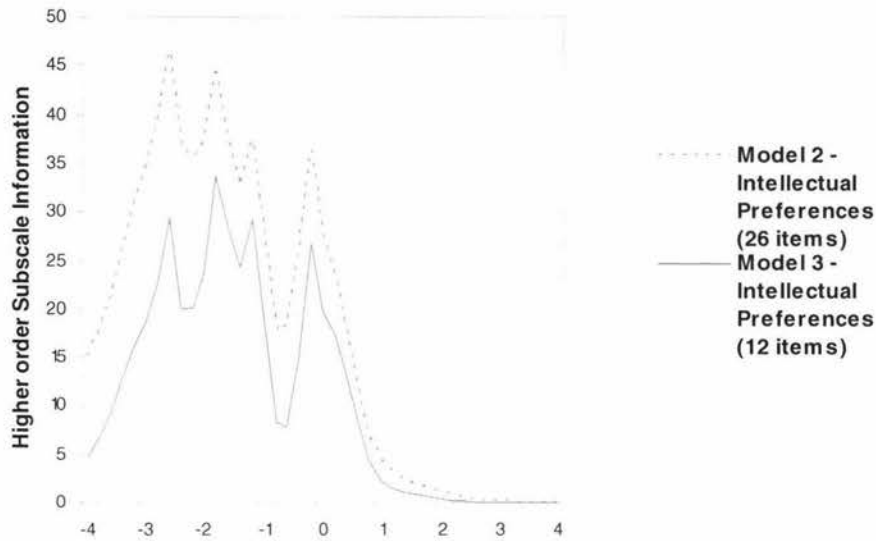
The distance between the two lines in the graph shown above illustrates the slight information reduction after discarding 14 of the 26 items in this second-order factor. The 12 items selected for this second-order factor for *Model 3* clearly accounts for the majority of the information captured using all 28 items. This outcome was repeated with each second-order factor (see Fig. 15).

Fig. 15

Comparison of the SIF for Four of the Second-order Factors between Model 2 and Model 3

Fig. 15 *(continued)*

Model 2 -
Interpersonal Style
(28 items)

Model 3 -
Interpersonal Style
(12 items)

Model 2 -
Self Management
and Drive
(68 items)

Model 3 -
Self Management
and Drive
(27 items)
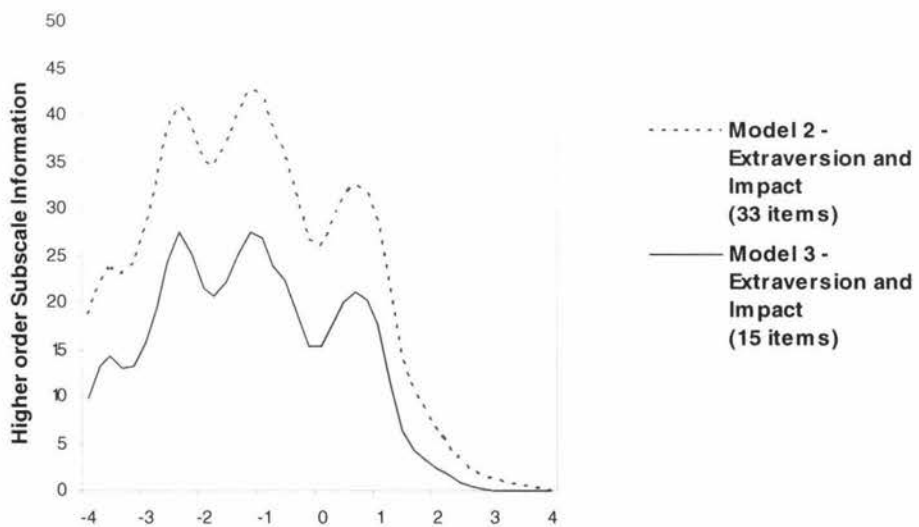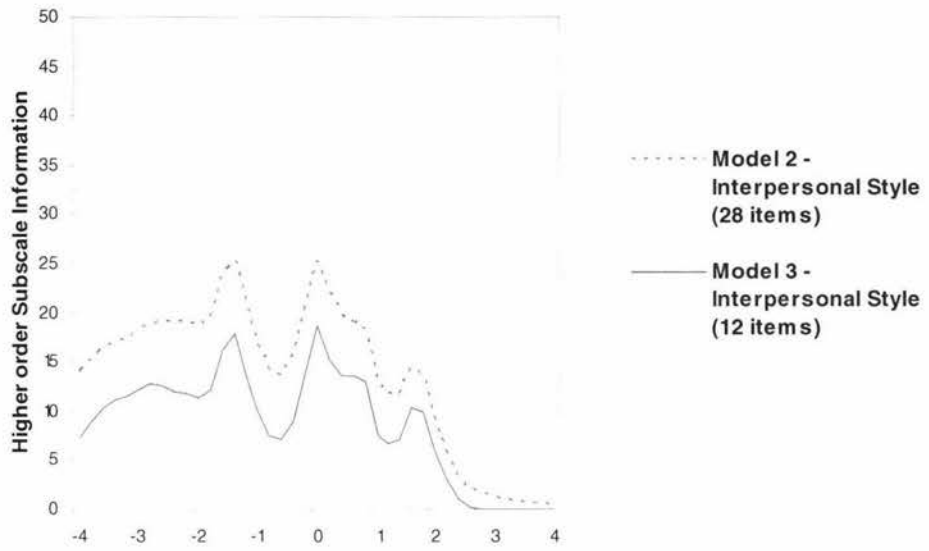
Fig. 15 *(continued)*
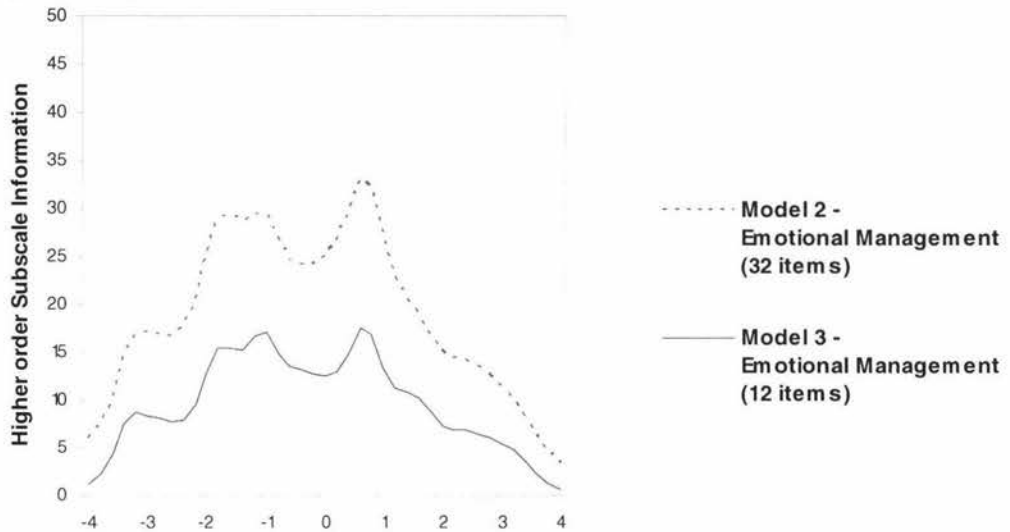
This is mainly due to the selection of items that showed higher 'a' parameters as shown in Table 11.

Table 11

Mean 'a' and 'b' Parameters for the Second-order Factors for Model 2 and Model 3

Second-order Factor	Model	M or SD	a	b1	b2	b3	b4
Extraversion and Impact	*Model 2*	**M**	1.77	-3.36	-1.74	-0.76	1.13
		SD	0.56	1.07	0.98	0.84	0.75
	Model 3	**M**	2.21	-2.94	-1.57	-0.74	0.84
		SD	0.53	0.73	0.74	0.67	0.73
Intellectual preferences	*Model 2*	**M**	1.85	-4.40	-2.68	-1.63	0.34
		SD	0.99	2.13	1.00	0.88	0.73
	Model 3	**M**	2.31	-3.44	-2.20	-1.22	0.45
		SD	1.24	1.01	0.69	0.41	0.47
Interpersonal Style	*Model 2*	**M**	1.45	-4.83	-2.24	-0.93	1.26
		SD	0.85	2.35	1.56	1.40	1.33
	Model 3	**M**	2.06	-4.40	-2.21	-1.17	0.48
		SD	0.94	1.98	1.45	1.40	1.03
Self Management and Drive	*Model 2*	**M**	1.57	-4.46	-2.72	-1.70	0.48
		SD	0.60	1.46	1.10	0.97	1.06
	Model 3	**M**	1.97	-3.78	-2.36	-1.44	0.40
		SD	0.52	0.92	0.96	0.83	0.71
Emotional Management	*Model 2*	**M**	1.52	-2.70	-0.65	0.36	2.67
		SD	0.58	2.47	1.80	1.55	1.65
	Model 3	**M**	1.94	-2.00	-0.40	0.45	2.29
		SD	0.55	0.96	1.02	0.99	0.98

The five second-order factors were combined for the final CFA and along with the fit statistics are shown in Fig. 16.

Fig. 16

CFA and Fit Statistics for Model 3

Model	Items	χ^2	DF	Prob	GFI	TLI	CFI	RMSEA
Model 3	78	6393.05	2896	0.00	0.73	0.75	0.76	0.05

Personality

Extraversion & Impact: Social Ease, Gregariousness, Self-Disclosure, Social Confidence, Assertiveness

Intellectual Preferences: Curiosity, Breadth of Interest, Innovation, Variety-Seeking

Interpersonal Style: Tolerance, Affiliation, Empathy, Self-belief

Self Management & Drive: Conscientiousness, Activity Level, Achievement Striving, Self-Discipline, Optimism, Planned, Locus of Control, Compliance, Perfectionism

Emotional Management: Anxiety, Approval Seeking, Emotional Responsiveness, Emotional Expression

- 46 -

In comparison to *Model 1* and *Model 2*, *Model 3* has greatly improved model fit due primarily to the selection of higher quality, higher information items, although in some cases the item wording was quite similar. Fit statistics for the three stages of item selection are shown in Table 12.

Table 12

Fit statistics for all Three Models to Illustrate the Effects of the Item Selection Process

Model	Items	χ^2	DF	Prob	GFI	TLI	CFI	RMSEA
Model 1								
	225	49436.32	24948	0.00	0.47	0.48	0.48	0.05
Model 2								
	187	33632.72	17174	0.00	0.53	0.57	0.58	0.05
Model 3								
	78	6393.05	2896	0.00	0.73	0.75	0.76	0.05

The GFI, TLI, and CFI have improved throughout this selection process. *Model 3* did not meet the preferred levels for these three fit statistics of .90 or better, however it has shown significant improvement over the original state of the questionnaire. The RMSEA has remained at good levels throughout this process indicating that the factor structure was supported and that the aim of reducing the items in this questionnaire without losing model fit has been successful.

The final assessment of this selection process is achieved by calculating a *TIF* in order to compare the total information captured by *Model 3* (78 items) in comparison to *Model 2* (187 items), as illustrated in Fig. 17. As was desired, all 26 first-order factors are represented in both of these models.

Fig. 17

Comparison of the TIF between Model 2 and Model 3
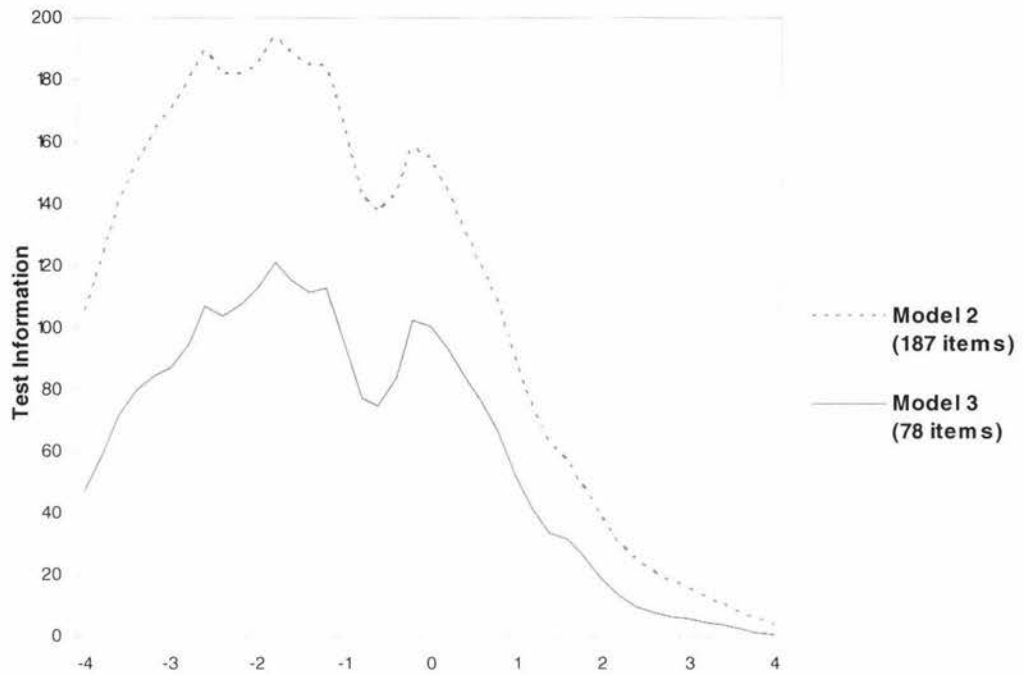
This *Test Information Function* clearly illustrates that *Model 3* captures the vast majority of the information captured by *Model 2* with far less input. This was achieved with only 78 items, as the average item discrimination in *Model 3* was superior to that of the items in *Model 2* (see Table 13).

Table 13

Comparison of Mean 'a' and 'b' Parameters between Model 2 and Model 3

Model	M or SD	a	b1	b2	b3	b4
Model 2	*M*	1.61	-3.98	-2.04	-0.96	1.15
	SD	0.73	2.07	1.53	1.38	1.40
Model 3	*M*	2.08	-3.39	-1.86	-0.94	0.79
	SD	0.74	1.33	1.19	1.08	1.01

Discussion

Objectives

The objectives of this research were to reduce the length of a personality questionnaire using *Confirmatory Factor Analysis* and *Item Response Theory*. The specific hypotheses were (1) *the questionnaire length could be reduced*, (2) *the model fit would improve*, and (3) *the reduction in information captured would be minimal relative to the item reduction*. All three hypotheses have been clearly satisfied illustrating the power of *Item Response Theory* for developing personality measures. These findings also confirm the proposition that there was a significant amount of item redundancy in this questionnaire, as there is hypothesised to be in many personality questionnaires in current use. This selection process can therefore be repeated with any other questionnaire and it is proposed that similar results to this study will unfold. This process has clearly increased the efficiency of this questionnaire, both for the participants and the interpreters of the results. As stated by Costsa and McCrae (1997) and Gerbing and Tuley (1991) test development is a continuous process, therefore this questionnaire with its reduced length and enhanced measurement precision can be used as an empirical advancement of the original model as well as the core of a future revision of this questionnaire.

In regards to the specific items that have been identified through this analysis, all users of Goldberg's online resource can now apply the findings of this study to their current or future personality models.

Some interesting aspects of the data did unfold through this analysis and therefore the specifics of the results follow. However, before drawing conclusions it is important to recognise the strengths of the analyses that were performed.

Strengths of this Analysis

In order to successfully specify models through CFA the size of the sample needs to increase with the amount of parameters the model is expected to estimate (Tabachnick & Fidell, 2007). The largest model in this analysis, and hence the model with the most parameters, was a 3-stage higher-order model with 225 individual items.

The required sample size increases further according to the amount of item response options (in this case five). This is due to the large amount of possible response combinations. Another factor affecting the sample size was the decision to split the dataset in two so that a different sample could be used for the CFA before the IRT analysis (n = 484) and the CFA after (n = 489), thus providing an important cross-validation of the model structure. In order to ensure maximum precision for the IRT analysis the full sample (n = 973) was used. Despite the combination of all of these factors the sample size was adequate for all tests in this analysis as was illustrated through the acceptable RMSEA levels throughout this development process (MacCallum, Browne, & Sugawara, 1996). By following these procedures the validity and strength of the results in this study was increased.

Stages of Development

Cross-loaded Items

The original state of this model had many items that cross-loaded. This is a common occurrence with personality testing as items are often not sufficiently specific and thus load on more than one factor. Although many of the cross-loaded items in this questionnaire could have potentially been shown to belong to only one factor and may have been incorrectly categorised under more than one factor, this process was not carried out. Perhaps the Thurstone method as discussed by Andrich (1996) may have assisted in this phase of the process.

Confirmatory Factor Analyses

The first CFA of the first-order factors showed that many factors had very low fit statistics and did not satisfy the principles of unidimensionality. When these were combined into a three-stage higher-order model, *Model 1*, the RMSEA was at a good level, however the other fit statistics were far lower than the preferred minimum levels. The SMC analyses of the first-order factors showed that many of the items did not highly relate to the other items in their factor. This also adversely affected the reliability of these factors. After the SMC deletions the reliability and fit statistics improved considerably for the majority of the first-order factors.

Two of the first-order factors failed to converge however reintroducing the previously discarded cross-loaded items rectified this. Additionally, as these items were deleted from their other first-order factors they no longer cross-loaded. Interestingly, item Q315 that was reintroduced (for *Achievement Striving*) was selected through the IRT analysis for *Model 3*. The SMC also showed item Q315 as a central item for this first-order factor and therefore its reintroduction was highly successful. The other first-order factor that originally failed to converge was *Conscientiousness*, as 9 out of its original 11 items cross-loaded. This implies that this first-order factor may need reworking to ascertain exactly what it is measuring and how it differs from the other first-order factors where its items also loaded.

Model 2 (187 items) showed greater fit than *Model 1* (225 items) and a reasonable state of unidimensionality was satisfied, thus the IRT analysis was permissible (Hambleton & Jones, 1993). The items selected through the IRT analysis had the greatest affect on the fit of the CFA model as the GFI, TFI, and CFI all improved by over .15 between *Model 2* (187 items) and *Model 3* (78 items). Importantly, the RMSEA remained at a good level of .05 throughout this process.

In all cases the improvement of model fit at the second-order level led to improved model fit at the third-order level. This result supports the hypothesised structure of the *Five-Factor Model*. Chernyshenko et al. (2001a) made a similar inference in regards to their analysis of the 16PF.

The loading of the second-order factors onto the third-order factor was also of interest with strong positive loadings for four of the five second-order factors (*Extraversion and Impact* .85, *Intellectual Preferences* .85, *Interpersonal Style* .70, and *Self Management and Drive* .85). *Emotional Management* was the only second-order factor showing a negative loading (-.50). This may be due to the meaning of the first-order factors, as someone who is high in *Anxiety* may have less control over their emotions and hence this may negatively affect *Personality*. However, in regards to the *Intellectual Preferences* second-order factor, someone who is high in *Curiosity* may have greater intellectual stimulus and this may therefore positively affect *Personality*. The individual items are best understood through analysing the results of the IRT selection process.

Item Response Theory

A total of 187 items were used in the IRT analysis. This resulted in 65 *poor items* (35%), 31 *mediocre items* (16%), and 91 *good items* (49%). *Model 3* had 78 items, the vast majority of which were from the *good item* category. The selections of these items were based on principles discussed earlier however some selections were simpler than others. In most cases the first one or two items selected for each first-order factor were straight forward as they were clearly the best items. The third item was often more difficult either because there were two *good* options and it was difficult to select the better of the two, or there were two *mediocre* options both of which were lacking quality in regard to one or more of the principles discussed earlier. Additionally, if the best available items were similar, one was replaced with the next best option. Occasionally there was no next best option and the similar items were selected. Specific examples illustrate the decisions that were made during this selection process.

The *Variety-seeking* and *Self-belief* first-order factors both had three good items. However they also both had two items that were similarly worded and hence captured very similar information. These items that are slight variations of each other are termed 'bloated specifics' (Cattell, 1978) and should be avoided where possible. Not only is repeating the same or similar questions a redundant exercise but it can also lead to mulitcollinearity (Fletcher & Hattie, 2004). Unfortunately, in the case of these two first-order factors, there were no other options and the similarly worded items were selected.

Item location and item wording was not assessed in cases where there were only three items available with good shape and good information. This was the case for seven of the 26 first-order factors. For example, three of the six *Affiliation* items in the IRT analysis were categorised as *good* and three as *poor*. Although in this case each selected item did contribute some unique information, in regards to other first-order factors the selection was forced. In cases where more than three options were available for selection the fourth and fifth items were also often worded similarly. In some cases the only difference was in regard to the item being positively or negatively worded, hence essentially asking the same question. Approximately five of the *Model 3* first-order factors included similarly worded items.

In every item selected for *Model 3* each response option provided some unique information, however in a small number of cases the neutral response (*neither agree nor disagree*) provided less than desired. This was because this response option was not used as much as the other options for these items. The preference is for all the peaks of the curves of the OCF to be reasonably high and of approximate equal height. However, 61 (33%) of the *Model 2* items would be categorised as 'low neutral response' items. It is often the case that the neutral response option is used less for certain types of question and in some cases this response option is completely dropped for revised editions of the measure (Kim & Hong, 2004; Hong et al., 2005). Alternatively, the distinction between items that did and did not suit the original Likert-scale could be identified and recommendations for the non-suited items could be made (Fletcher & Hattie, 2004). In this case a 4-point Likert-scale may improve the measurement precision for these items. This may then alleviate the deletion of many of the items and lead to a better final selection of items in terms of wording variety.

Occasionally an item captured less information than would have been preferred, however the shape and location was satisfactory and the item was selected. This occurred with the *Breadth of Interest* first-order factor, as only four items were available for the IRT analysis so there were few to select from. The low number of items in this first-order factor was due to the need for the items to be reasonably specific so as to not capture the same information as the *Variety-seeking* or *Curiosity* first-order factors, with which three of its original items cross-loaded. In order to identify the best items for these three first-order factors, their semantic distinctions need to be improved.

The 'a' parameter can be used as an important indicator for the quality of an item, as through assessing the relative graphs items with higher 'a' parameters showed better shape and location than those with lower 'a' parameters. This is illustrated through the average 'a' parameter increase between *Model 2* ($a = 1.61$) and *Model 3* ($a = 2.08$). However, as stated, 33% of the *Model 2* items had a low neutral response option yet many still had reasonably high 'a' parameters. Therefore, it is important and necessary to visually assess the item graphs, as visual principles such as (1) *response option use* and (2) *variation of item information* cannot be inferred from the numerical 'a' parameter on its own.

The other important feature of the graphs is the invaluable information gained through assessing the TIF as was performed by Gomez et al. (2005). The image of the questionnaire shown through this graph can exhibit its strengths and weaknesses as well as any reduction of information from reducing the total items in a measure. The overall measurement properties of this questionnaire would be improved if some of the items were reworded or new items added in order to capture information from the positive end of the personality continuum. This would provide a more balanced questionnaire and capture important information that is currently being missed. The current balance can be understood in terms of the purpose of the questionnaire in that it is attempting to identify potential issues within personality and the majority of issues may occur at the negative end of the personality continuum. In regards to the comparison of information captured between *Model 2* and *Model 3*, it is clearly shown that *Model 3* (78 items) with less than half of the items of *Model 2* (187 items) has captured the majority of the information. This image communicates more to the viewer than any words or numbers in this analysis and clearly illustrates the possibilities available through IRT.

Reliability

Although not a focus of this analysis, reliability was recorded for the first-order factors at each stage of this process. These comparisons showed that on average reliability only decreased by .06 for each first-order factor between *Model 2* and *Model 3*. This was a minimal difference considering that the *Model 2* first-order factors had on average more than seven items and all the *Model 3* first-order factors had only three items. Similarly worded items in each first-order factor would have achieved even higher reliability, however where possible the goal was to capture information under the whole trait.

Lie Scale

A further and final aspect for discussion is the importance of the lie scale, *Desirable Virtues*. As this scale is not part of the factor structure for this questionnaire, it could theoretically be comprised of items from the other first-order factors. Lie scale items tend to include the words '*never*' or '*always*' in order to capture extreme beliefs which when framed positively seem desirable. By including the analysis of this scale in the appendices of this study two tasks were achieved. Firstly, an example of how a single factor can be analysed using the procedures described in this study is demonstrated.

Secondly, according to Ferrando and Chico (2001), although IRT procedures are strong and have the possibility to detect the response patterns of participants, such as faking a good response, they state that IRT is not as strong as the inclusion of a psychometrically proven *Lie* or *Social Desirability* scale. In contrast to this, Costa and McCrae (1997) indicated that the validity and utility of lie scales has not been proven and therefore they did not include a lie scale in the revised *NEO-PI*. Nonetheless, the *Desirable Virtues* analysis has been included in this study and individual researchers can decide for themselves whether or not its use is valid.

Limitations

The main limitation faced in this study was the number of items available for IRT selection. Hong et al. (2005) applied IRT to a scale of 180 items, identifying 25 for the final scale (14% selected). In the current study the original first-order factors were usually 10 items in length and three of these were selected for the final model (30% selected). These selections were predominantly *good* items however as has been stated some of the first-order factors have room for improvement. This also meant that eight of the cross-loaded items had to be re-introduced in order for two of the first-order factors to survive the first stage of analysis. This was not preferable however the results justified the means as the final model had three quality items for each of the original 26 first-order factors, thus satisfying one of the aims of this study. Additionally, the statistical significance of any study is greater with increased sample size. The current sample size was adequate however larger samples would have been even better, especially in regards to the first CFA, which included 225 items.

Practical Implications

This study illustrates an empirical example of the application of *Item Response Theory* to the development of a personality questionnaire. Through this analysis, not only have the benefits of IRT been explained and proven, but also the grounds for the progression from *Classical Test Theory* to *Item Response Theory* have been unquestionably demonstrated. This analysis has shown good validity through both the application of unidimensionality analyses and the use of a cross-validation sample. Each lower-order factor was shown to be psychometrically valid and through this the hypothesised structure of the *Big-Five* was confirmed as appropriate for this data.

Stringent procedures were followed which can be replicated by others and no better solution was seen to be available for these items. The final model included some items that captured similar information and some that were possibly better suited to a 4-point Likert-scale, however the overall model was shown to fit well through the final CFA.

The organisation that provided the data for this analysis is able to apply these findings to all future uses of their questionnaire, saving time whilst also producing more accurate results. They will have the knowledge about which items actually provide quality information and which are psychometrically redundant. The importance of certain personality characteristics for organisational outcomes (i.e. Barrick et al., 2001) enhances the need for the identification of items that best express these characteristics, and that is precisely what has been achieved through this analysis.

IRT is a statistically complex procedure that can consume great amounts of time (McKinley, 1989), however due to the underlying principles of IRT, other users of this particular item bank can now apply the findings of this study to their research. Therefore, this analysis has many practical implications for individuals and organisations such as a better understanding of which items to select for past and future uses of these items and an enhanced starting point for the further development of this questionnaire. The fundamentals of IRT also mean that these items are sample independent and that the responses to these items can be assessed at the vastly more practical item level, thus greatly enhancing their utility.

Future Considerations
The first future research consideration is in regard to the Likert-scale used in this analysis. Over 33% of the items from *Model 2* showed a low neutral response option. Many of these items could have potentially been selected for the final model as the other response options showed good shape and good location however due the unused response option they were discarded. A future analysis of these items where participants respond on a 4-point scale may show improved measurement precision. A decision would have to be made whether to change all the items, as was performed by Kim and Hong (2004) and Hong et al. (2005), or whether to only change the items that showed this pattern, as was recommended by Fletcher and Hattie (2004).

Fletcher and Hattie (2004) also found significant differences between the information contributed by the negatively worded items and the positively worded items in their analysis of the PSDQ. For this reason it would be of interest to perform the same analysis for the items in this measure as the results may indicate areas of potential improvement for both the total information captured by this questionnaire and the location of the information.

One further aspect of importance is the concurrent validity of the factors in this measure. This topic was not focused on in this study however it is prudent to confirm whether these items relate to other pre-existing measures of personality and whether the interpretations of the factors are satisfactory.

As stated by Costa and McCrae (1997), test development is an ongoing process. Therefore, the final suggestion for future research is the continuation of the process that has begun in this study. This could be achieved through: an analysis of each individual first-order factor in regards to their face validity; changing the Likert-scale for each item as appropriate; some of the first-order factors need items added to achieve an improved spread of information; and finally some of the first-order factors may benefit from having items replaced, especially in regards to similarly worded items. This process would serve as an excellent revision of this questionnaire and would once again greatly improve its measurement precision for future use.

Conclusions

All hypotheses were met in this study. The benefits of *Item Response Theory* have been illustrated and a reduced length questionnaire has been empirically identified. Due to the sample independence of analyses performed with *Item Response Theory* any person can use the results of this study, and the individual item responses can be legitimately interpreted at the item level. The complete process used through this study is a clear example of how any researcher can greatly enhance the measurement precision of any personality questionnaire they seek to investigate or develop. Other researchers are therefore encouraged to move away from *Classical Test Theory* and towards the psychometric benefits that are available for test development through *Item Response Theory*.

References

Andrich, D. (1996). A hyperbolic cosine latent trait model for unfolding polytomous responses: Reconciling Thurstone and Likert methodologies. *British Journal of Mathematical and Statistical Psychology.* 49, 347–365.

Arbuckle, J. L. (1999). AMOS 4.0: Analysis of moment structures. [Computer Program]. Chicago: Smallwaters.

Baker, F. (2001). The Basics of Item Response Theory. ERIC Clearinghouse on Assessment and Evaluation, University of Maryland, College Park, MD. Available Online: http://edres.org/irt/baker

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology.* 44, 1-25.

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). The FFM personality dimensions and Job performance: Meta-Analysis of Meta-Analyses. Invited submission to a special "selection" issues of *International Journal of Selection and Assessment.* 9, 9-30.

Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin.* 107, 238–246.

Bollen, K.A. (1989). A new incremental fit index for general structural equation models. *Sociological Methods and Research.* 17, 303–316.

Bolt, D. M., Hare, R. D., Vitale, J. E., & Newman, J. P. (2004). A Multigroup Item Response Theory Analysis of the Psychopathy Checklist—Revised. *Psychological Assessment.* 16, 155–168.

Borkenau, P. & Ostendorf, F. (1990). Comparing exploratory and confirmatory factor analysis: A study on the 5-factor model of personality. *Personality and Individual Differences.* 11, 515-524.

Browne, M. W., & Cudeck, R. (1993). Alternate ways of assessing model fit. In K. A. Nollem & J. S. Long (Eds.). Testing structural equation models. (136-162). Newbury Park, CA: Sage.

Byrne, B. M. (2001). Structural Equation Modelling With AMOS: Basic Concepts, Applications, and Programming (Multivariate Applications Series) (Paperback)

Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. *Journal of Abnormal and Social Psychology.* 38, 476-506.

Cattell, R. B. (1978). The scientific use of factor analysis in the behavioral and life sciences. New York: Plenum Press.

Chernyshenko, O. S., Stark, S., & Chan, K. (2001a). Investigating the hierarchical factor structure of the Fifth Edition of the 16PF: An application of the Schmid-Leiman orthogonalization procedure. *Educational and Psychological Measurement.* 61, 290-302.

Chernyshenko, O. S., Stark, S., Chan, K., Drasgow, F., & Williams B. (2001b). Fitting Item Response Theory Models to Two Personality Inventories: Issues and Insights. *Multivariate Behavioral Research.* 36, 523-562.

Churchill, G. A., Jr. (1979). A Paradigm for Developing Better Measures of Marketing Constructs. *Journal of Marketing Research.* 16, 64-73.

Conn, S. R., & Rieke, M. L. (Eds.). (1994). The 16PF fifth edition technical manual. Champaign, IL: Institute for Personality and Ability Testing.

Costa Jr. P. T, & McCrae, R. R. (1997). Stability and Change in Personality Assessment: The Revised NEO Personality Inventory in the Year 2000. *Journal of Personality Assessment.* 68, 86-94.

De Ayala, R.J *(1993).* An introduction to polytomous item response theory models. *Measurement & Evaluation in Counseling & Development.* 25, 748-1756.

Eysenck, H. J. (1992). For ways five factors are not basic. *Personality and Individual Differences.* 13, 667-673.

Ferrando, P. J. & Chico, E. (2001). Detecting dissimulation in personality test scores: A comparison between person-fit indices and detection scales. *Educational and Psychological Measurement.* 61, 997-1012.

Fletcher, R. B., & Hattie, J. A. (2004). An examination of the psychometric properties of the physical self-description questionnaire using a polytomous item response model. *Psychology of Sport and Exercise*. 5, 423–446

Fletcher, R. B., & Hattie, J. A. (2005). Gender differences in physical self-concept: a multidimensional differential item functioning analysis. *Educational and Psychological Measurement*. 65, 657-667.

Gefen, D. (2003). Assessing unidimensionality through lisrel: an explanation and example. *Communications of the Association for Information Systems*. 12, 23-47.

Gerbing, D. W., & Tuley, M. R. (1991). The 16PF related to the fine-factor model of personality: multiple-indicator measurement versus the a priori scales. *Multivariate Behavioural Research*. 26, 271-89.

Goldberg, L. R. (1990). An alternative "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology*. 59, 1216-1229.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. C. (2006). The International Personality Item Pool and the future of public-domain personality measures. *Journal of Research in Personality*. 40, 84-96.

Gomez, R., Cooper, A., & Gomez, A. (2005). An item response theory analysis of the Carver and White (1994) BIS/BAS Scales. *Personality and Individual Differences*. 39, 1093–1103.

Guenole, N., & Chernyshenko, O. S. (2005). The Suitability of Goldberg's Big Five Personality Markers in New Zealand: A Dimensionality, Bias, and Criterion Validity Evaluation. *New Zealand Journal of Psychology*. 34, 85-95.

Hambleton, R. K., Jones, R. W. (1993). Comparison of classical test theory and item response theory and their applications to test development. *Educational Measurement: Issues and Practice*. 12, 253-262.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). Fundamentals of item response theory. London: Sage.

Hattie, J. A. (1984). An empirical study of various indices for determining unidimensionality. *Multivariate Behavioral Research.* 19, 49-78.

Hattie, J. A. (1985). Methodology review: Assessing unidimensionality of tests and items. *Applied Psychological Measurement.* 9, 139-164.

Hong, S., Kim, B. S. K., & Wolfe, M. M. (2005). A Psychometric Revision of the European American Values Scale for Asian Americans Using the Rasch Model. *Measurement & Evaluation in Counseling & Development.* 37, 194-207.

International Personality Item Pool: A Scientific Collaboratory for the Development of Advanced Measures of Personality Traits and Other Individual Differences (http://ipip.ori.org/). Internet Web Site.

Kim, B. S. K., & Hong, S. (2004). A Psychometric Revision of the Asian Values Scale Using the Rasch Model. *Measurement & Evaluation in Counseling & Development.* 37, 15-27.

MacCallum, R. C., Browne, M. W., Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods.* 1, 130-149.

McKinley, R. L. (1989). METHODS, PLAINLY SPEAKING: An Introduction to Item Response Theory. *Measurement and Evaluation in Counselling and Development.* 22, 37-57.

Mislevy, R. J. & Bock, R. D. (1991). *BILOG* user's guide. [Computer Program]. Chicago, IL: Scientific Software.

Ones, D. S., & Viswesvaran, C. (1998). Gender, Age, and Race Differences on Overt Integrity Tests: Results Across Four Large-Scale Job Applicant Data Sets. *Journal of Applied Psychology.* 83, 35-42.

Ostini, R. Nering, M. L. (2006). Polytomous Item Response Theory Models. Sage Publications. Thousand Oaks.

Ozer, D. J. & Reise, S. P. (1994). Personality assessment. *Annual Review of Psychology*. 45, 357-389.

Raju, N. S., Laffitte, L. J., & Byrne, B. M. (2002). Measurement Equivalence: A Comparison of Methods Based on Confirmatory Factor Analysis and Item Response Theory. *Journal of Applied Psychology*. 87, 517–529.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement*. 34, 100.

SPSS for Windows, Rel. 13.0. (2004). [Computer Program]. Chicago: SPSS Inc

Steiger, J. H. & Lind, J. M. (1980). Statistically based tests for the number of common factors. Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.

Tabachnick, B. G. & Fidell, L. S. (2007). Using multivariate statistics. Boston : Pearson/Allyn & Bacon.

Tanaka, J. S., & Huba, G. J. (1984). Structures of Psychological Distress: Testing confirmatory hierarchical models. *Journal of Consulting and Clinical Psychology*. 52, 719-721.

Thissen, D. (1991). *MULTILOG* user's guide: Multiple, categorical item analysis and test scoring using item response theory (Version 6.0). [Computer Program]. Chicago, IL: Scientific Software.

Tuerlinckx, F., De Boeck, P., & Lens, W. (2002). Measuring Needs With the Thematic Apperception Test: A Psychometric Study. *Journal of Personality and Social Psychology*. Vol. 82, No. 3, 448–461.

Wang, W., Chen, P., & Cheng, Y. (2004). Improving Measurement Precision of Test Batteries Using Multidimensional Item Response Models. *Psychological Methods*. 9, 116–136.

Zickar, M. J. (1998). Modelling Item-level data with Item Response Theory. *Current directions in Psychological Science*. 7, 104-109.

Appendices

Appendix 1

Original list of Questionnaire Items

Second-order Factor	First-order Factor	Item	Pos or Neg	Wording
Extraversion and Impact				
	Social Ease			
		Q005	+	Make friends easily.
		Q231	+	Warm up quickly to others.
		Q039	+	Feel comfortable around people.
		Q070	+	Act comfortably with others.
		Q044	+	Cheer people up.
		Q066	−	Am hard to get to know.
		Q263	−	Often feel uncomfortable around others.
		Q085	−	Avoid contacts with others.
		Q030	−	Am not really interested in others.
		Q302	−	Keep others at a distance.
	Gregariousness			
		Q296	+	Love large parties.
		Q123	+	Talk to a lot of different people at parties.
		Q009	+	Enjoy being part of a group.
		Q245	+	Involve others in what I am doing.
		Q167	+	Love surprise parties.
		Q055	−	Prefer to be alone.
		Q178	−	Want to be left alone.
		Q041	−	Don't like crowded events.
		Q257	−	Avoid crowds.
		Q187	−	Seek quiet.
	Self-Disclosure			
		Q274	+	Am open about my feelings.
		Q036	+	Am open about myself to others.
		Q307	+	Disclose my intimate thoughts.
		Q305	+	Show my feelings.
		Q059	+	Talk about my worries.
		Q243	−	Don't talk a lot.
		Q043	−	Reveal little about myself.
		Q283	−	Bottle up my feelings.
		Q011	−	Have little to say.
		Q242	−	Say little.
	Social-Confidence			
		Q039	+	Feel comfortable around people.
		Q118	+	Don't mind being the centre of attention.
		Q020	+	Am good at making impromptu speeches.
		Q151	+	Express myself easily.
		Q065	+	Have a natural talent for influencing people.
		Q189	−	Hate being the centre of attention.
		Q053	−	Lack the talent for influencing people.
		Q263	−	Often feel uncomfortable around others.
		Q247	−	Don't like to draw attention to myself.
		Q011	−	Have little to say.
	Assertiveness			
		Q198	+	Take charge.
		Q050	+	Try to lead others.
		Q091	+	Can talk others into doing things.
		Q008	+	Seek to influence others.
		Q191	+	Take control of things.
		Q297	−	Wait for others to lead the way.
		Q289	−	Keep in the background.
		Q011	−	Have little to say.
		Q247	−	Don't like to draw attention to myself.
		Q188	−	Hold back my opinions.

Second-order Factor	First-order Factor	Item	Pos or Neg	Wording
Intellectual Preferences				
	Curiosity			
		Q180	+	Like to solve complex problems.
		Q086	+	Love to read challenging material.
		Q268	+	Have a rich vocabulary.
		Q135	+	Can handle a lot of information.
		Q046	+	Enjoy thinking about things.
		Q240	−	Am not interested in abstract ideas.
		Q018	−	Avoid philosophical discussions.
		Q260	−	Have difficulty understanding abstract ideas.
		Q083	−	Am not interested in theoretical discussions.
		Q301	−	Avoid difficult reading material.
	Breadth of Interest			
		Q086	+	Love to read challenging material.
		Q254	+	Find political discussions interesting.
		Q120	+	Am interested in many things.
		Q143	+	Love to learn new things.
		Q212	+	Want to increase my knowledge.
		Q110	−	Do not enjoy going to art museums.
		Q261	−	Do not like art.
		Q111	−	Prefer to stick with things that I know.
		Q301	−	Avoid difficult reading material.
		Q064	−	Do not like concerts.
	Innovation			
		Q278	+	Love to think up new ways of doing things.
		Q166	+	Have a vivid imagination.
		Q100	+	Am full of ideas.
		Q295	+	Carry the conversation to a higher level.
		Q228	+	Come up with bold plans.
		Q199	+	Have excellent ideas.
		Q318	−	Do not have a good imagination.
		Q310	−	Have difficulty imagining things.
		Q165	−	Will not probe deeply into a subject.
		Q094	−	Can't come up with new ideas.
	Variety-Seeking			
		Q210	+	Prefer variety to routine.
		Q278	+	Love to think up new ways of doing things.
		Q112	+	Am open to change.
		Q133	+	Enjoy hearing new ideas.
		Q010	+	Seek adventure.
		Q220	+	Like to begin new things.
		Q172	+	Like to visit new places.
		Q063	−	Don't like the idea of change.
		Q294	−	Dislike changes.
		Q111	−	Prefer to stick with things that I know.

Second-order Factor	First-order Factor	Item	Pos or Neg	Wording
Interpersonal Style				
	Tolerance			
		Q125	+	Accept people as they are.
		Q224	+	Respect others.
		Q275	+	Sympathise with the homeless.
		Q234	+	Believe there are many sides to most issues.
		Q142	+	Believe that others have good intentions.
		Q104	−	Am a bad loser.
		Q072	−	Get irritated easily.
		Q164	−	Lay down the law to others.
		Q129	−	Treat people as inferiors.
		Q312	−	Am quick to judge others.
		Q073	−	Am annoyed by others' mistakes.
	Afflilation			
		Q103	+	Am easy to satisfy.
		Q057	+	Can't stand confrontations.
		Q159	+	Hate to seem pushy.
		Q075	−	Have a sharp tongue.
		Q040	−	Contradict others.
		Q250	−	Love a good fight.
		Q290	−	Yell at people.
		Q185	−	Insult people.
		Q298	−	Get back at others.
		Q121	−	Hold a grudge.
	Empathy			
		Q090	+	Feel others' emotions.
		Q013	+	Anticipate the needs of others.
		Q099	+	Reassure others.
		Q096	+	Make others feel good.
		Q280	+	Am concerned about others.
		Q221	+	Have a good word for everyone.
		Q116	+	Make people feel welcome.
		Q236	+	Take time out for others.
			−	(No negative items.)
	Self-belief			
		Q262	+	Dislike being the center of attention.
		Q087	+	Dislike talking about myself.
		Q006	+	Consider myself an average person.
		Q128	+	Seldom toot my own horn.
		Q024	−	Believe that I am better than others.
		Q175	−	Think highly of myself.
		Q137	−	Have a high opinion of myself.
		Q311	−	Know the answers to many questions.
		Q299	−	Boast about my virtues.
		Q107	−	Make myself the center of attention.

Second-order Factor	First-order Factor	Item	Pos or Neg	Wording
Self Management and Drive				
	Conscientiousness			
		Q147	+	Am always prepared.
		Q309	+	Pay attention to details.
		Q080	+	Get chores done right away.
		Q259	+	Carry out my plans.
		Q277	+	Make plans and stick to them.
		Q033	−	Waste my time.
		Q233	−	Find it difficult to get down to work.
		Q315	−	Do just enough work to get by.
		Q157	−	Don't see things through.
		Q251	−	Shirk my duties.
	Activity Level			
		Q130	+	Am always busy.
		Q049	+	Am always on the go.
		Q127	+	Do a lot in my spare time.
		Q223	+	Can manage many things at the same time.
		Q273	+	React quickly.
		Q060	−	Like to take it easy.
		Q266	−	Like to take my time.
		Q146	−	Like a leisurely lifestyle.
		Q252	−	Let things proceed at their own pace.
		Q056	−	React slowly.
	Achievement-Striving			
		Q218	+	Go straight for the goal.
		Q182	+	Work hard.
		Q183	+	Turn plans into actions.
		Q079	+	Plunge into tasks with all my heart.
		Q229	+	Do more than what's expected of me.
		Q015	+	Set high standards for myself and others.
		Q149	+	Demand quality.
		Q027	−	Am not highly motivated to succeed.
		Q315	−	Do just enough work to get by.
		Q207	−	Put little time and effort into my work.
	Self-Discipline			
		Q080	+	Get chores done right away.
		Q147	+	Am always prepared.
		Q161	+	Start tasks right away.
		Q256	+	Get to work at once.
		Q259	+	Carry out my plans.
		Q233	−	Find it difficult to get down to work.
		Q033	−	Waste my time.
		Q152	−	Need a push to get started.
		Q156	−	Have difficulty starting tasks.
		Q306	−	Postpone decisions.
	Optimism			
		Q077	+	Never give up hope.
		Q281	+	Love life.
		Q276	+	Work on improving myself.
		Q197	+	Keep improving myself.
		Q003	+	Know what I want.
		Q140	−	Feel that my life lacks direction.
		Q113	−	Am not sure where my life is going.
		Q200	−	Am resigned to my fate.
		Q150	−	Let others determine my choices.
		Q047	−	Agree to anything.

Appendix 1 *(continued)*

Second-order Factor	First-order Factor	Item	Pos or Neg	Wording
	Planned			
		Q001	+	Like to plan ahead.
		Q211	+	Do things by the book.
		Q217	+	Am exacting in my work.
		Q309	+	Pay attention to details.
		Q277	+	Make plans and stick to them.
		Q195	–	Jump into things without thinking.
		Q054	–	Like to act on a whim.
		Q163	–	Often make last-minute plans.
		Q264	–	Make rash decisions.
		Q201	–	Make a mess of things.
	Locus of Control			
		Q208	+	Feel comfortable with myself.
		Q232	+	Believe that my success depends on ability rather than luck.
		Q267	+	Just know that I will be a success.
		Q022	+	Come up with good solutions.
		Q281	+	Love life.
		Q070	+	Act comfortably with others.
		Q106	+	Feel up to any task.
		Q249	+	Like to take responsibility for making decisions.
		Q155	+	Take the initiative.
		Q179	+	Make a decision and move on.
		Q214	–	Believe that unfortunate events occur because of bad luck.
		Q216	–	Believe that the world is controlled by a few powerful people.
		Q140	–	Feel that my life lacks direction.
		Q031	–	See difficulties everywhere.
		Q225	–	Habitually blow my chances.
		Q045	–	Believe that some people are born lucky.
		Q093	–	Dislike taking responsibility for making decisions.
		Q265	–	Am less capable than most people.
		Q244	–	Dislike myself.
		Q316	–	Feel that I'm unable to deal with things.
	Compliance			
		Q105	+	Try to follow the rules.
		Q170	+	Keep my promises.
		Q288	+	Pay my bills on time.
		Q205	+	Tell the truth.
		Q004	+	Listen to my conscience.
		Q061	–	Break rules.
		Q048	–	Break my promises.
		Q193	–	Get others to do my duties.
		Q088	–	Do the opposite of what is asked.
		Q317	–	Misrepresent the facts.
	Perfectionism			
		Q269	+	Continue until everything is perfect.
		Q271	+	Want every detail taken care of.
		Q023	+	Want everything to be "just right."
		Q141	+	Want things to proceed according to plan.
		Q230	+	Demand perfection in others.
		Q204	+	Keep a sharp eye on others' work.
		Q168	+	Expect dedicated work from others.
		Q203	–	Am not bothered by messy people.
		Q238	–	Am not bothered by disorder.

Appendix 1 *(continued)*

Second-order Factor	First-order Factor	Item	Pos or Neg	Wording
Emotional Management				
	Anxiety			
		Q206	+	Get stressed out easily.
		Q177	+	Worry about things.
		Q026	+	Get upset easily.
		Q021	+	Have frequent mood swings.
		Q089	+	Often feel blue.
		Q139	−	Am relaxed most of the time.
		Q032	−	Am not easily bothered by things.
		Q109	−	Rarely get irritated.
		Q303	−	Seldom feel blue.
		Q145	−	Am not easily frustrated.
	Approval Seeking			
		Q016	+	Worry about what people think of me.
		Q122	+	Conform to others' opinions.
		Q062	+	Need the approval of others.
		Q222	+	Want to amount to something special in others' eyes.
		Q174	+	Do what others do.
		Q253	−	Don't care what others think.
		Q078	−	Am not concerned with making a good impression.
		Q235	−	Feel it's OK that some people don't like me.
		Q162	−	Want to form my own opinions.
		Q082	−	Want to be different from others.
	Emotional Responsiveness			
		Q131	+	Think about the causes of my emotions.
		Q144	+	Pay a lot of attention to my feelings.
		Q304	+	Am usually aware of the way that I'm feeling.
		Q171	+	Notice my emotions.
		Q215	+	Often stop to analyse how I'm feeling.
		Q196	−	Rarely think about how I feel.
		Q052	−	Rarely analyse my emotions.
		Q282	−	Am not in touch with my feelings.
		Q158	−	Often ignore my feelings.
		Q270	−	Rarely notice my emotional reactions.
	Emotional Expression			
		Q246	+	Experience my emotions intensely.
		Q090	+	Feel others' emotions.
		Q124	+	Am passionate about causes.
		Q114	+	Enjoy examining myself and my life.
		Q126	+	Try to understand myself.
		Q258	−	Seldom get emotional.
		Q190	−	Am not easily affected by my emotions.
		Q270	−	Rarely notice my emotional reactions.
		Q176	−	Experience very few emotional highs and lows.
		Q007	−	Don't understand people who get emotional.

Appendix 2

Items from Original Questionnaire that Cross-loaded

Cross-loaded Item	Wording	First-order Factor Where Item Loaded
Q011	Have little to say.	*Self-Disclosure* *Social-Confidence* *Assertiveness*
Q033	Waste my time.	*Conscientiousness* *Self-Discipline*
Q039	Feel comfortable around people.	*Social Ease* *Social-Confidence*
Q070	Act comfortably with others.	*Social Ease* *Locus of Control*
Q080	Get chores done right away.	*Conscientiousness* *Self-Discipline*
Q086	Love to read challenging material.	*Curiosity* *Breadth of Interest*
Q090	Feel others' emotions.	*Empathy* *Emotional Expression*
Q111	Prefer to stick with things that I know.	*Breadth of Interest* *Variety-Seeking*
Q140	Feel that my life lacks direction.	*Optimism* *Locus of Control*
Q147	Am always prepared.	*Conscientiousness* *Self-Discipline*
Q233	Find it difficult to get down to work.	*Conscientiousness* *Self-Discipline*
Q247	Don't like to draw attention to myself.	*Social-Confidence* *Assertiveness*
Q259	Carry out my plans.	*Conscientiousness* *Self-Discipline*
Q263	Often feel uncomfortable around others.	*Social Ease* *Social-Confidence*
Q270	Rarely notice my emotional reactions.	*Emotional Responsiveness* *Emotional Expression*
Q277	Make plans and stick to them.	*Conscientiousness* *Planned*
Q278	Love to think up new ways of doing things.	*Innovation* *Variety-Seeking*
Q281	Love life.	*Optimism* *Locus of Control*
Q301	Avoid difficult reading material.	*Curiosity* *Breadth of Interest*
Q309	Pay attention to details.	*Conscientiousness* *Planned*
Q315	Do just enough work to get by.	*Conscientiousness* *Achievement-Striving*

Appendix 3

CFA for Second-order Factors for Model 1

Intellectual Preferences

Items	χ^2	DF	Prob	GFI	TLI	CFI	RMSEA
32	1748.09	460	0.00	0.79	0.73	0.75	0.08
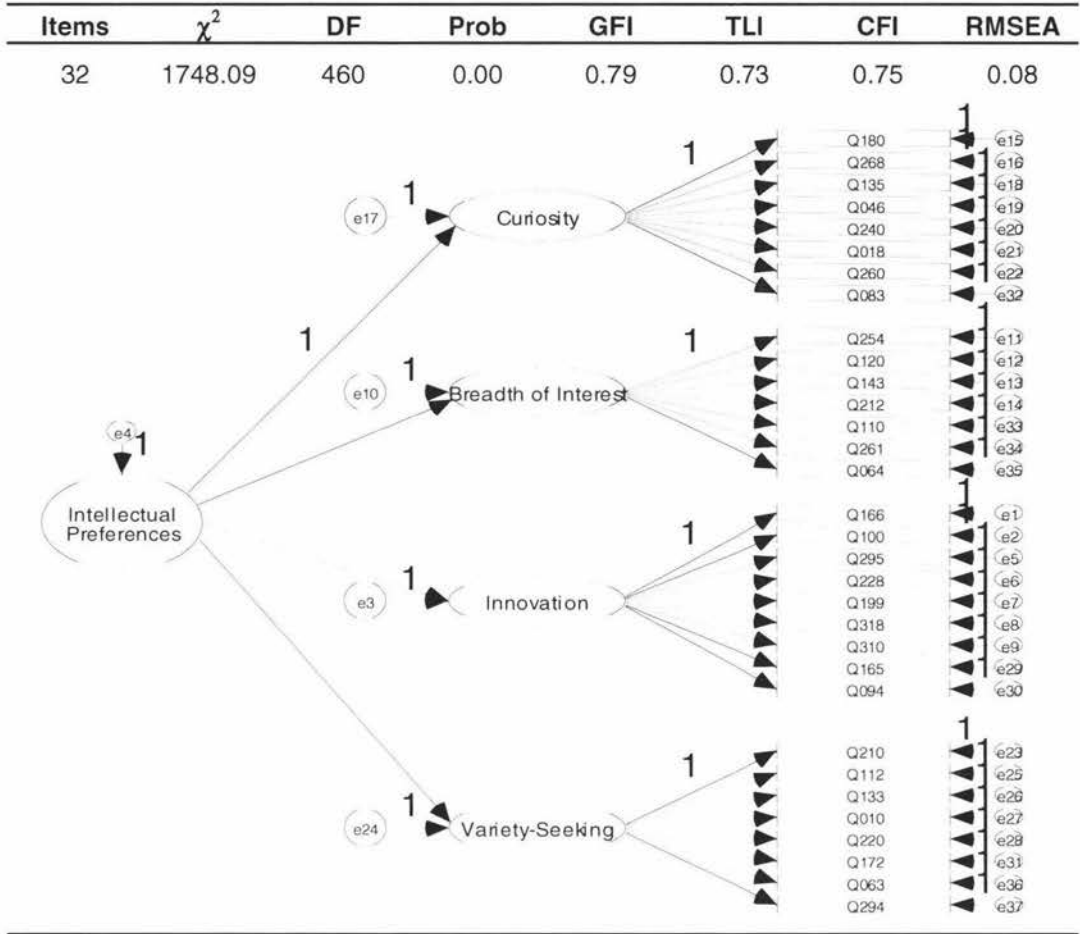
Appendix 3 *(continued)*

Interpersonal Style

Items	χ^2	DF	Prob	GFI	TLI	CFI	RMSEA
38	2085.83	662	0.00	0.79	0.64	0.66	0.07

Self Management and Drive

Items	χ^2	DF	Prob	GFI	TLI	CFI	RMSEA
78	6689.77	2916	0.00	0.70	0.64	0.65	0.05

Self Management & Drive

Conscientiousness — Q157, Q251

Activity Level — Q130, Q049, Q127, Q223, Q273, Q060, Q266, Q146, Q252, Q056

Achievement Striving — Q218, Q182, Q183, Q079, Q229, Q015, Q149, Q027, Q207

Self-Discipline — Q161, Q256, Q152, Q156, Q306

Optimism — Q276, Q197, Q003, Q113, Q200, Q150, Q047, Q077

Planned — Q001, Q211, Q217, Q195, Q054, Q163, Q264, Q201

Locus of Control — Q208, Q232, Q267, Q022, Q106, Q249, Q155, Q179, Q214, Q216, Q031, Q225, Q045, Q093, Q265, Q244, Q316

Compliance — Q105, Q170, Q288, Q205, Q004, Q061, Q048, Q193, Q088, Q317

Perfectionism — Q269, Q271, Q023, Q141, Q230, Q204, Q168, Q203, Q238

Emotional Management

Items	χ^2	DF	Prob	GFI	TLI	CFI	RMSEA
37	1905.73	626	0.00	0.79	0.73	0.75	0.07

Appendix 4

CFA for Second-order Factors for Model 2

Extraversion and Impact

Items	χ^2	DF	Prob	GFI	TLI	CFI	RMSEA
33	1716.52	491	0.00	0.81	0.78	0.80	0.07

Intellectual Preferences

Items	χ^2	DF	Prob	GFI	TLI	CFI	RMSEA
26	1144.79	295	0.00	0.83	0.78	0.80	0.08

Interpersonal Style

Items	χ^2	DF	Prob	GFI	TLI	CFI	RMSEA
28	1032.88	347	0.00	0.86	0.77	0.79	0.06

Self Management and Drive

Items	χ^2	DF	Prob	GFI	TLI	CFI	RMSEA
68	5292.42	2201	0.00	0.72	0.70	0.71	0.05

Appendix 4 *(continued)*

Emotional Management

Items	χ^2	DF	Prob	GFI	TLI	CFI	RMSEA
32	1335.18	461	0.00	0.84	0.79	0.81	0.06
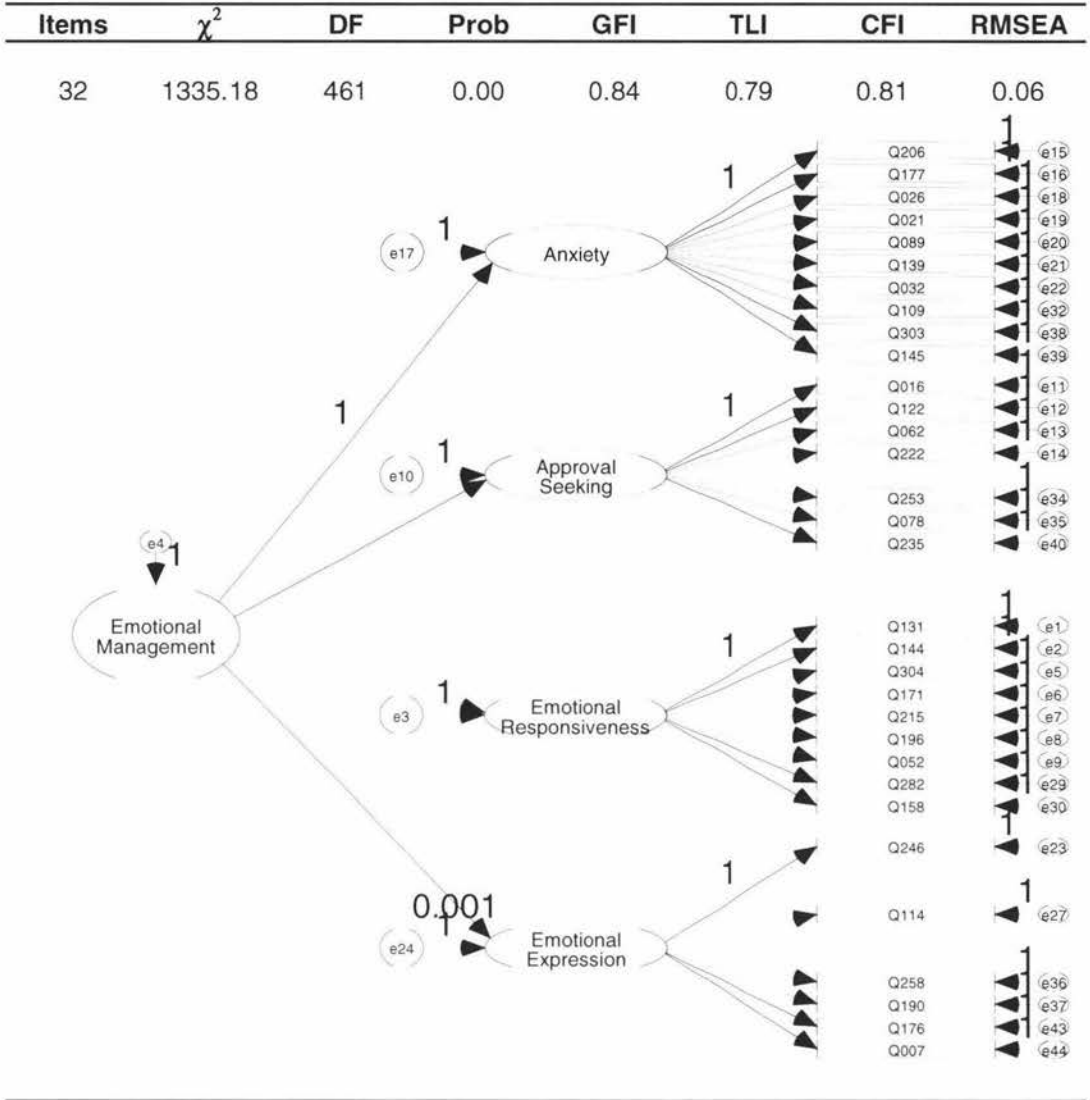
Appendix 5

CCC's, OCF's, and IIF's for the items in Model 3

Category Characteristic Curve	

Category Characteristic Curve

a = Item discrimination parameter

——□—— b1 = Intersection between option 1 & 2

— ◇ — b2 = Intersection between option 2 & 3

· · · △ · · · b3 = Intersection between option 3 & 4

× b4 = Intersection between option 4 & 5
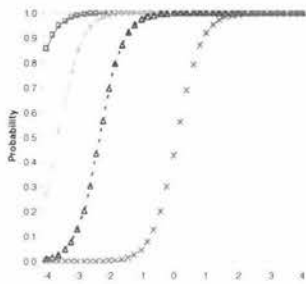
Operating Characteristic Function

——□—— 1- Very Inaccurate

— ◇ — 2- Moderately Inaccurate

3- Neither Inaccurate nor Accurate

4- Moderately Accurate

—— 5- Very Accurate

Item Information Function Curve

Extraversion & Impact   *Social Ease*

Q005	*Make friends easily.*			
a	b1	b2	b3	b4
2.14	-3.35	-1.97	-0.96	0.72
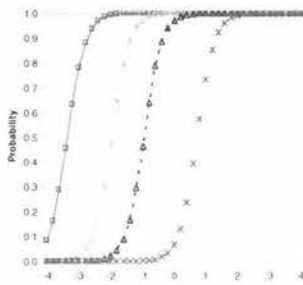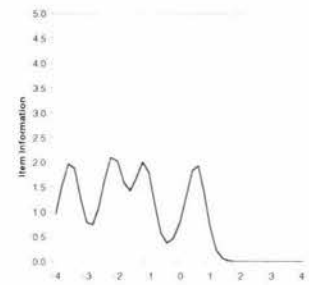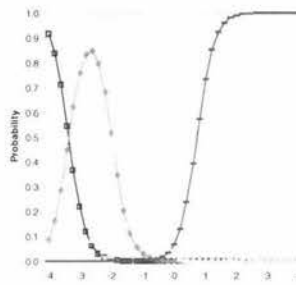
Extraversion & Impact   *Social Ease*

Q066	*Am hard to get to know.*			
a	b1	b2	b3	b4
1.99	-2.71	-1.11	-0.51	0.92
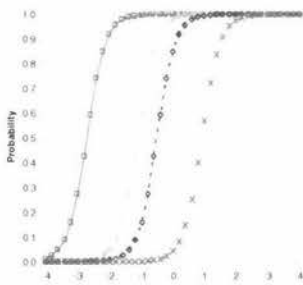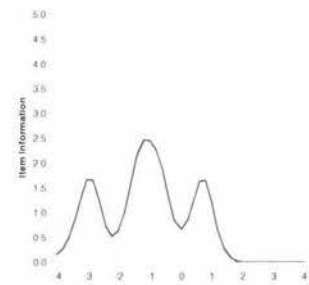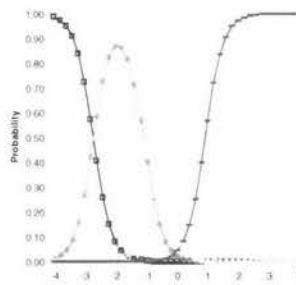
Extraversion & Impact   *Social Ease*

Q231	*Warm up quickly to others.*			
a	b1	b2	b3	b4
1.94	-3.91	-1.96	-0.93	1.19
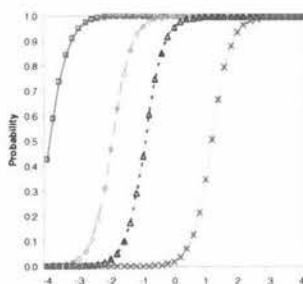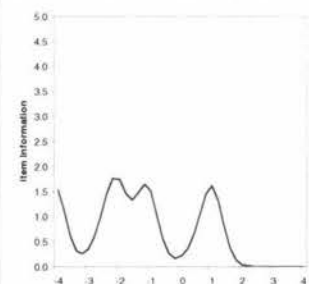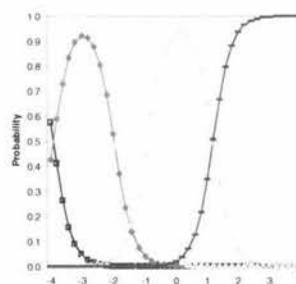
Appendix 5 *(continued)*

Category Characteristic Curve	*Operating Characteristic Function*	*Item Information Function Curve*

Category Characteristic Curve
- a = Item discrimination parameter
- b1 = Intersection between option 1 & 2
- b2 = Intersection between option 2 & 3
- b3 = Intersection between option 3 & 4
- b4 = Intersection between option 4 & 5

Operating Characteristic Function
- 1- Very Inaccurate
- 2- Moderately Inaccurate
- 3- Neither Inaccurate nor Accurate
- 4- Moderately Accurate
- 5- Very Accurate

Extraversion & Impact   *Gregariousness*

Q041 *Don't like crowded events.*

a	b1	b2	b3	b4
2.16	-1.9	-0.61	0.12	1.12
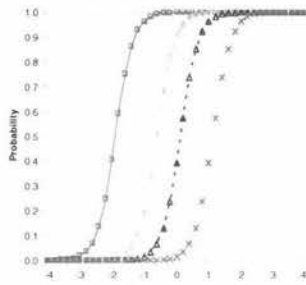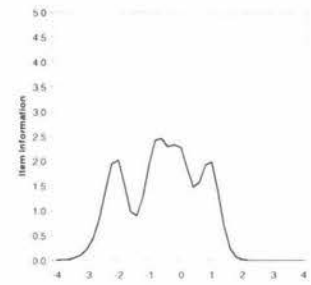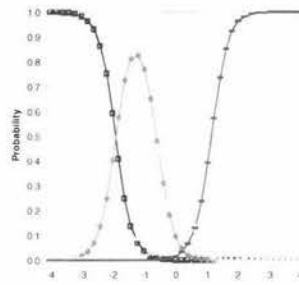
Extraversion & Impact   *Gregariousness*

Q257 *Avoid crowds.*

a	b1	b2	b3	b4
2.62	-2.36	-1.00	-0.32	0.92
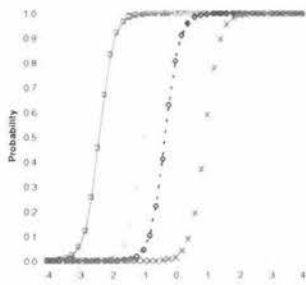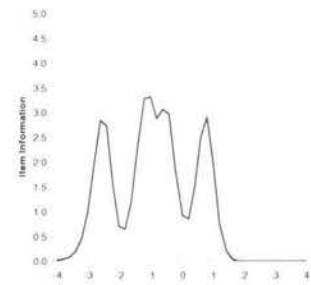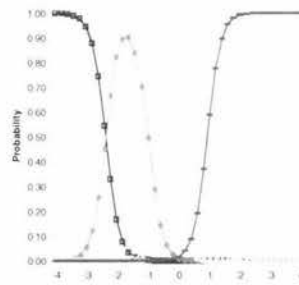
Extraversion & Impact   *Gregariousness*

Q296 *Love large parties.*

a	b1	b2	b3	b4
2.11	-1.77	-0.65	0.34	1.69
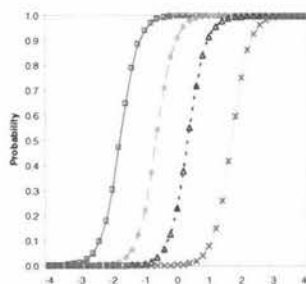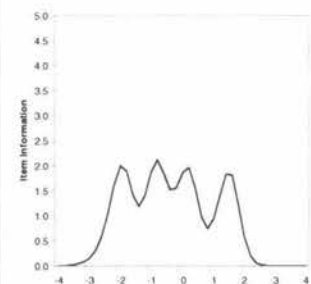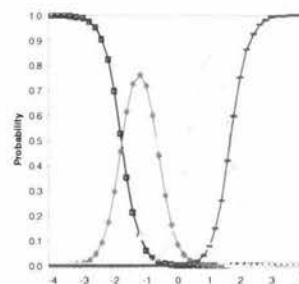
Appendix 5 *(continued)*

<u>Category Characteristic Curve</u>

a = Item discrimination parameter

———□——— b1 = Intersection between option 1 & 2

— — ◇ — — b2 = Intersection between option 2 & 3

· · · △ · · · b3 = Intersection between option 3 & 4

× b4 = Intersection between option 4 & 5

<u>Operating Characteristic Function</u>

———□——— 1- Very Inaccurate

———◇——— 2- Moderately Inaccurate

3- Neither Inaccurate nor Accurate

4- Moderately Accurate

——— 5- Very Accurate

<u>Item Information Function Curve</u>

Extraversion & Impact   *Self-Disclosure*

Q036 *Am open about myself to others.*

a	b1	b2	b3	b4
2.05	-2.64	-1.39	-0.76	0.84
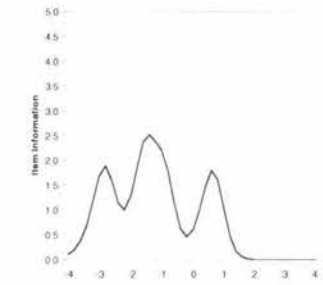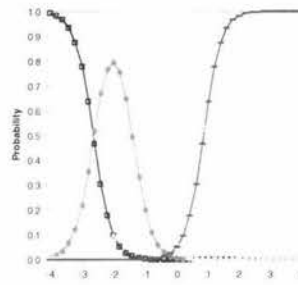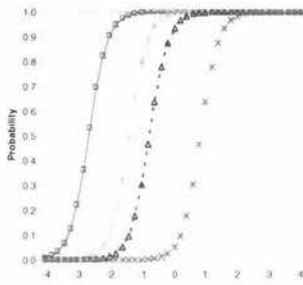
Extraversion & Impact   *Self-Disclosure*

Q274 *Am open about my feelings.*

a	b1	b2	b3	b4
3.46	-2.32	-1.00	-0.46	1.11
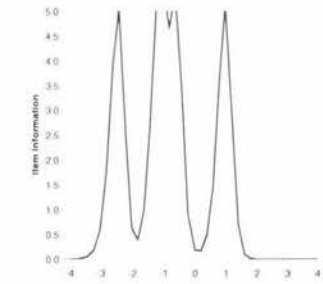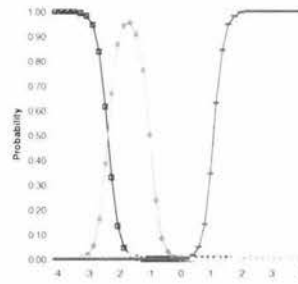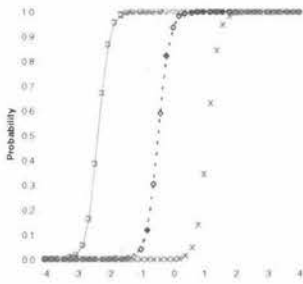
Extraversion & Impact   *Self-Disclosure*

Q305 *Show my feelings.*

a	b1	b2	b3	b4
1.89	-2.62	-0.68	0.10	2.22
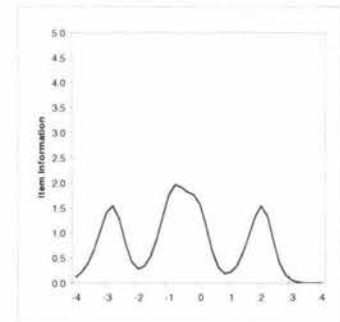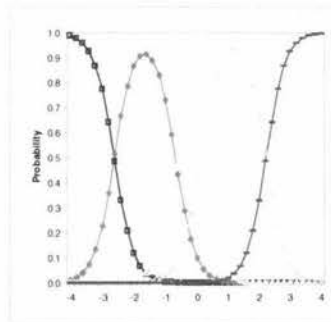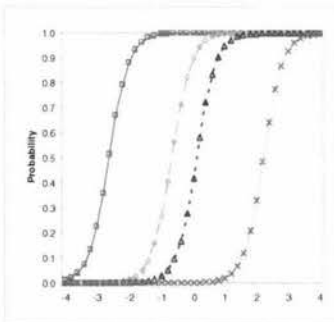
- 81 -

Appendix 5 *(continued)*

Item Information Function Curve

Extraversion & Impact   ***Social Confidence***

Q020   *Don't mind being the centre of attention.*

a	b1	b2	b3	b4
1.41	-2.46	-1.18	-0.4	1.33
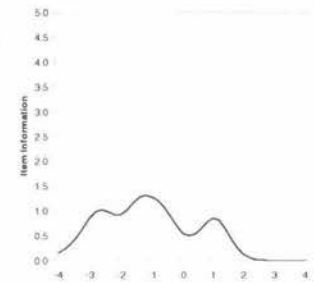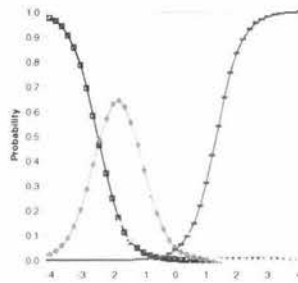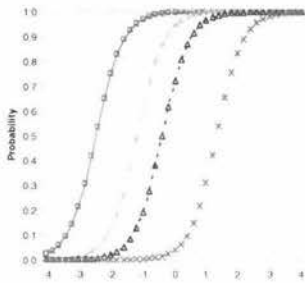
Extraversion & Impact   ***Social Confidence***

Q053   *Am good at making impromptu speeches.*

a	b1	b2	b3	b4
1.83	-3.52	-2.58	-1.83	0.00
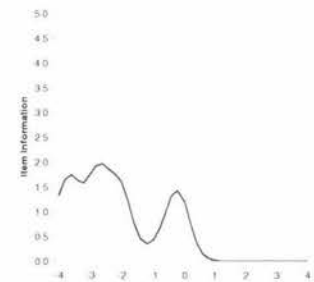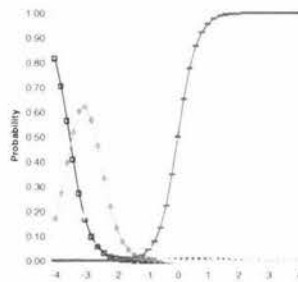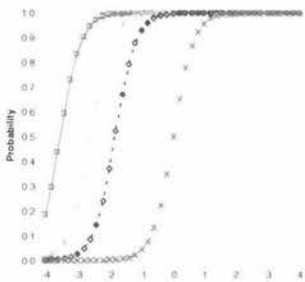
Extraversion & Impact   ***Social Confidence***

Q065   *Express myself easily.*

a	b1	b2	b3	b4
2.00	-3.24	-1.96	-0.91	0.99
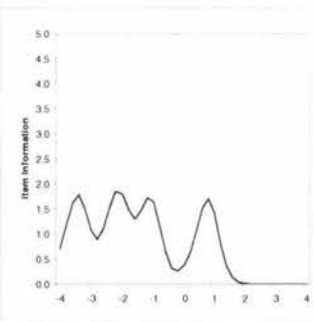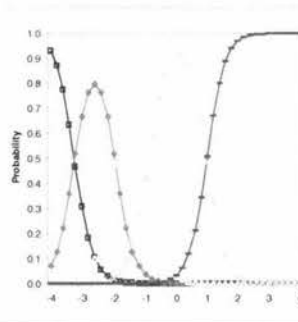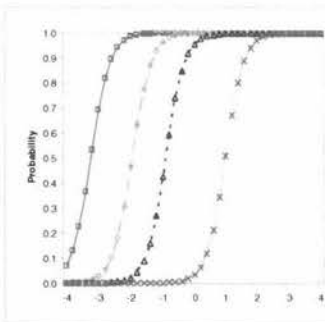
Appendix 5 *(continued)*

Category Characteristic Curve	*Operating Characteristic Function*	*Item Information Function Curve*

a = Item discrimination parameter
—⊟— **b1** = Intersection between option 1 & 2
———— **b2** = Intersection between option 2 & 3
- - -△- - - **b3** = Intersection between option 3 & 4
× **b4** = Intersection between option 4 & 5

—⊟— 1- Very Inaccurate
—◇— 2- Moderately Inaccurate
3- Neither Inaccurate nor Accurate
4- Moderately Accurate
———— 5- Very Accurate

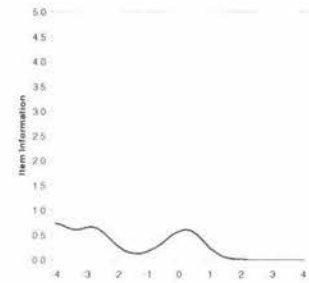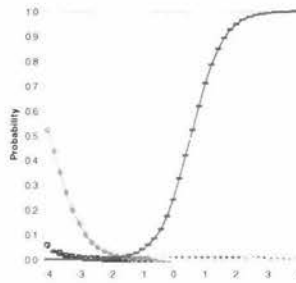## Extraversion & Impact   *Assertiveness*

Q050   *Try to lead others.*

a	b1	b2	b3	b4
2.08	-3.45	-2.7	-1.6	0.18
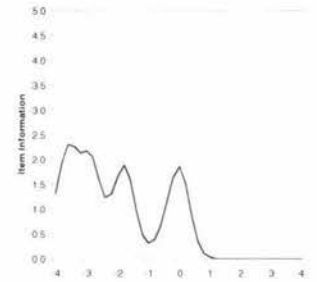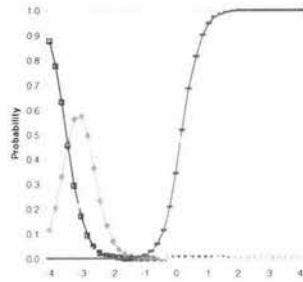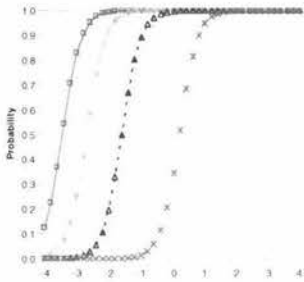
Extraversion & Impact   *Assertiveness*

Q191   *Take control of things.*

a	b1	b2	b3	b4
2.16	-4.19	-2.58	-1.62	0.62
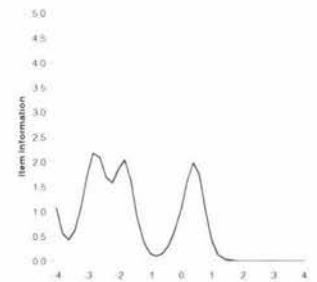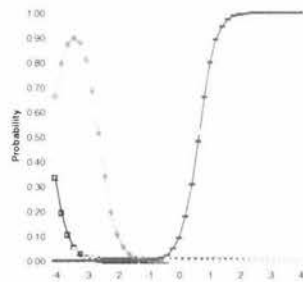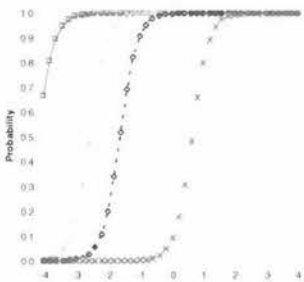
Extraversion & Impact   *Assertiveness*

Q198   *Take charge.*

a	b1	b2	b3	b4
3.24	-3.59	-2.17	-1.33	0.56

Appendix 5 *(continued)*
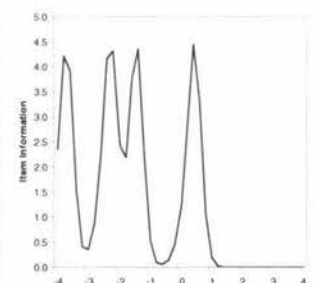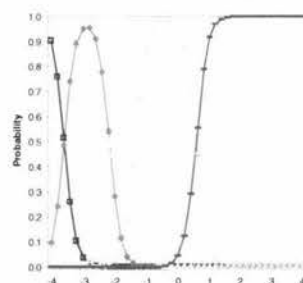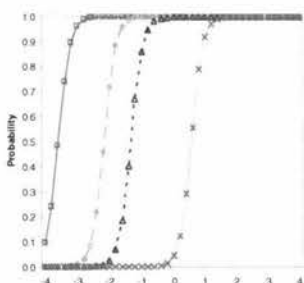
Intellectual Pref   *Curiosity*

Q018   *Avoid philosophical discussions.*

a	b1	b2	b3	b4
1.77	-3.07	-1.75	-0.86	0.55
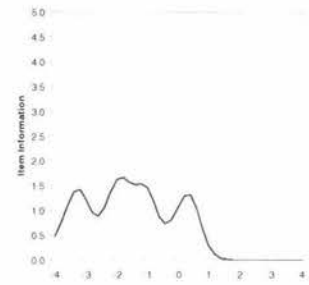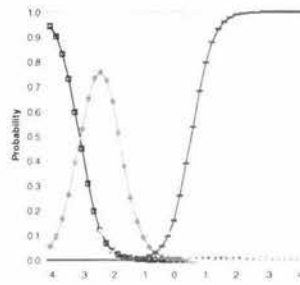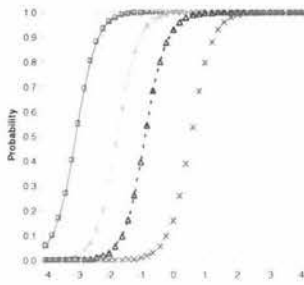
Intellectual Pref   *Curiosity*

Q083   *Am not interested in theoretical discussions.*

a	b1	b2	b3	b4
2.24	-2.91	-1.71	-1.08	0.43
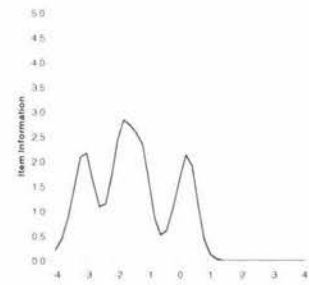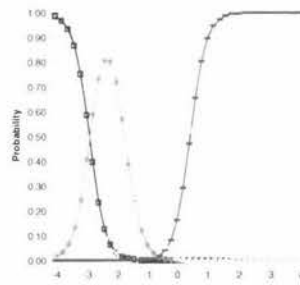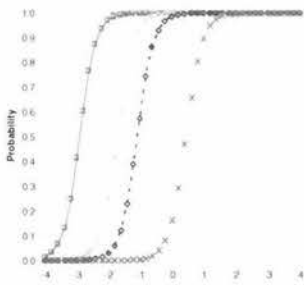
Intellectual Pref   *Curiosity*

Q240   *Am not interested in abstract ideas.*

a	b1	b2	b3	b4
2.66	-2.71	-1.79	-1.17	0.23
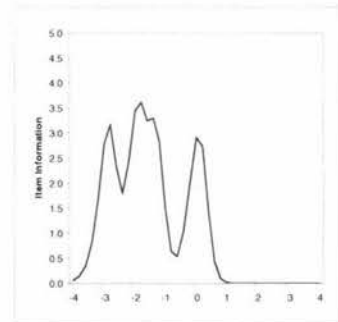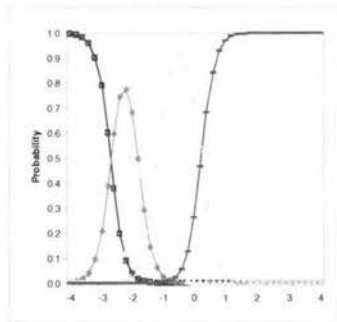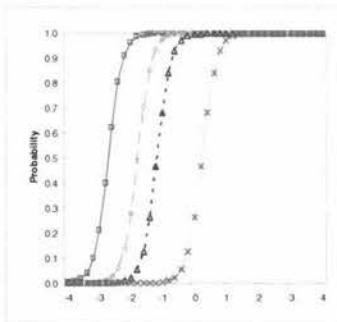
Appendix 5 *(continued)*

<u>Category Characteristic Curve</u>

- a = Item discrimination parameter
- —□— b1 = Intersection between option 1 & 2
- — ◇ — b2 = Intersection between option 2 & 3
- - - -△- - - b3 = Intersection between option 3 & 4
- ✕ b4 = Intersection between option 4 & 5

<u>Operating Characteristic Function</u>

- —□— 1- Very Inaccurate
- — ◇ — 2- Moderately Inaccurate
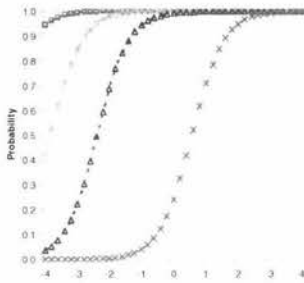- 3- Neither Inaccurate nor Accurate
- 4- Moderately Accurate
- —— 5- Very Accurate

<u>Item Information Function Curve</u>

Intellectual Pref   *Breadth of Interest*

Q064   *Do not like concerts.*

a	b1	b2	b3	b4
0.79	-5.56	-3.54	-1.89	0.24

Intellectual Pref   *Breadth of Interest*
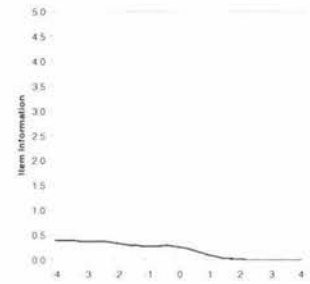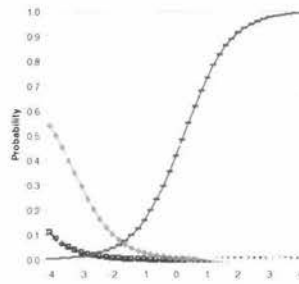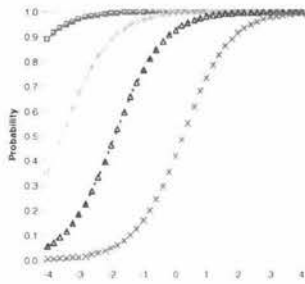
Q110   *Do not enjoy going to art museums.*

a	b1	b2	b3	b4
2.13	-2.23	-1.35	-0.76	0.33

Intellectual Pref   *Breadth of Interest*
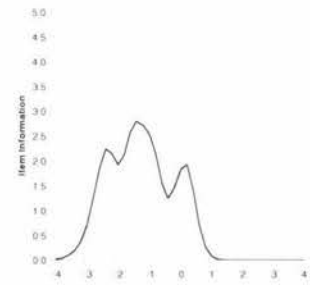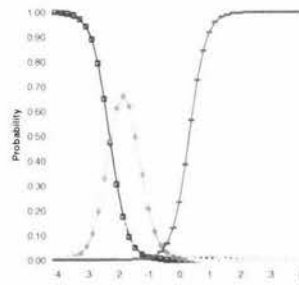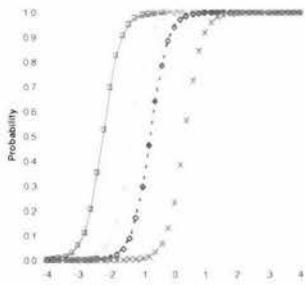
Q261   *Do not like art.*

a	b1	b2	b3	b4
5.79	-2.55	-1.68	-1.07	-0.15
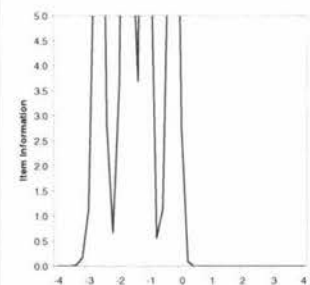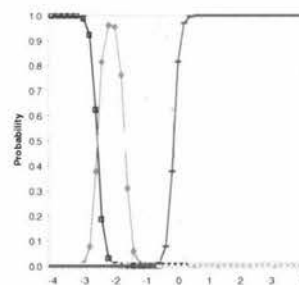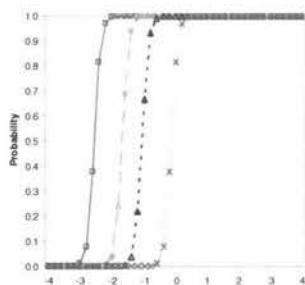
Appendix 5 *(continued)*

Category Characteristic Curve

a = Item discrimination parameter

⊟ **b1** = Intersection between option 1 & 2

▷ **b2** = Intersection between option 2 & 3

△ **b3** = Intersection between option 3 & 4

✕ **b4** = Intersection between option 4 & 5

Operating Characteristic Function

⊟ 1- Very Inaccurate

◇ 2- Moderately Inaccurate

3- Neither Inaccurate nor Accurate

4- Moderately Accurate

5- Very Accurate

Item Information Function Curve

Intellectual Pref   *Innovation*

Q100   *Am full of ideas.*

a	b1	b2	b3	b4
2.27	-3.39	-2.23	-1.21	0.46

Intellectual Pref   *Innovation*
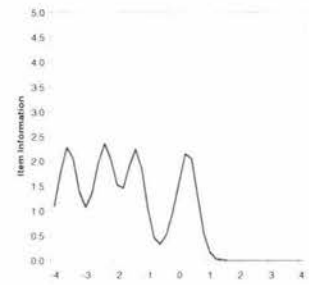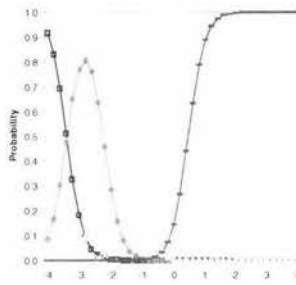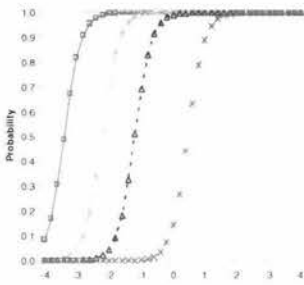
Q199   *Have excellent ideas.*

a	b1	b2	b3	b4
1.92	-4.45	-2.90	-1.32	0.75

Intellectual Pref   *Innovation*
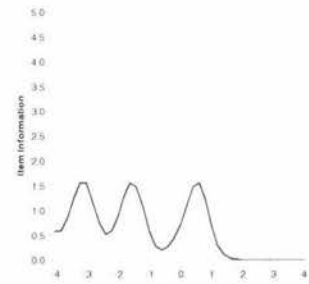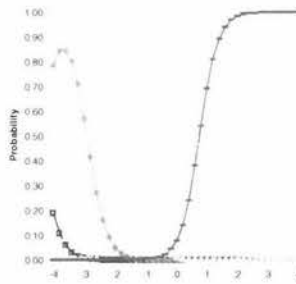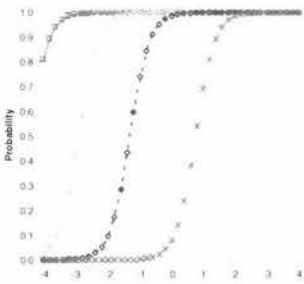
Q228   *Come up with bold plans.*

a	b1	b2	b3	b4
1.36	-3.65	-1.92	-0.60	1.74
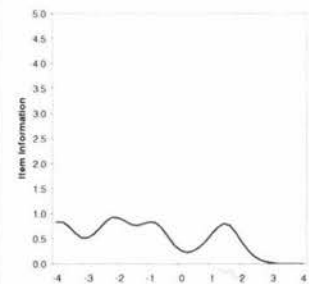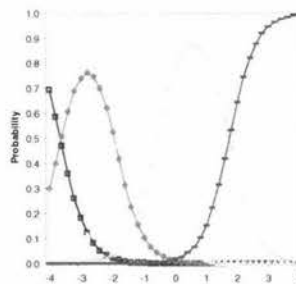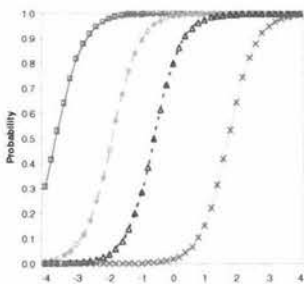
Appendix 5 *(continued)*

Intellectual Pref   *Variety-Seeking*

Q063   *Don't like the idea of change.*

a	b1	b2	b3	b4
2.55	-3.01	-2.09	-1.36	0.05
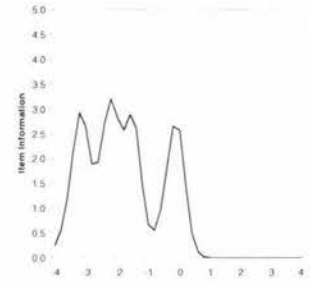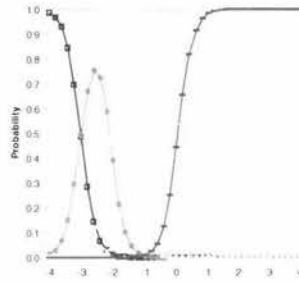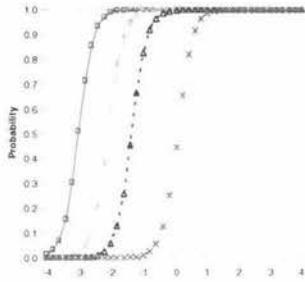
Intellectual Pref   *Variety-Seeking*

Q220   *Like to begin new things.*

a	b1	b2	b3	b4
1.48	-4.84	-3.31	-1.96	0.53
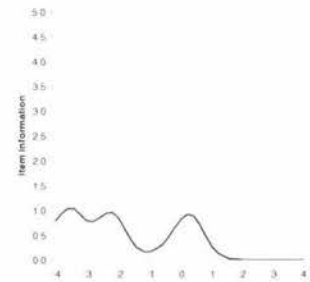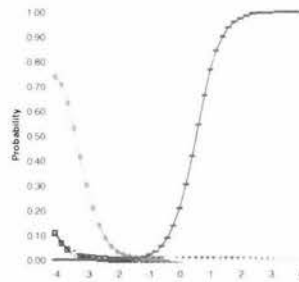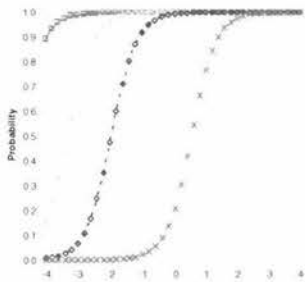
Intellectual Pref   *Variety-Seeking*

Q294   *Dislike changes.*

a	b1	b2	b3	b4
2.80	-2.86	-2.09	-1.41	0.22
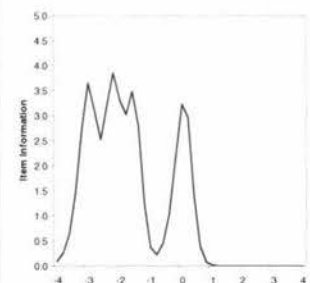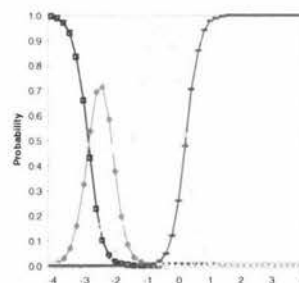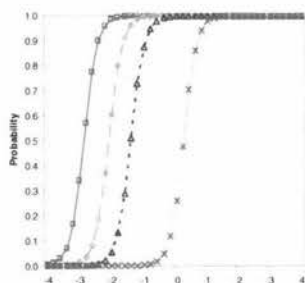
Appendix 5 *(continued)*

<u>Category Characteristic Curve</u>

a = Item discrimination parameter

—□— **b1** = Intersection between option 1 & 2

- -◇- - **b2** = Intersection between option 2 & 3

- - -△- - - **b3** = Intersection between option 3 & 4

× **b4** = Intersection between option 4 & 5

<u>Operating Characteristic Function</u>

—□— 1- Very Inaccurate

—◇— 2- Moderately Inaccurate

3- Neither Inaccurate nor Accurate

4- Moderately Accurate

——— 5- Very Accurate

<u>Item Information Function Curve</u>

**Interpersonal Style**   *Tolerance*

Q125   *Accept people as they are.*

a	b1	b2	b3	b4
1.23	-6.19	-2.74	-1.82	0.63
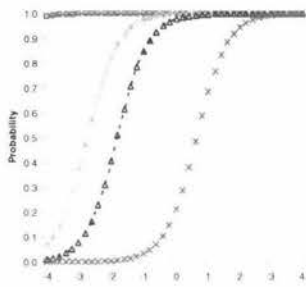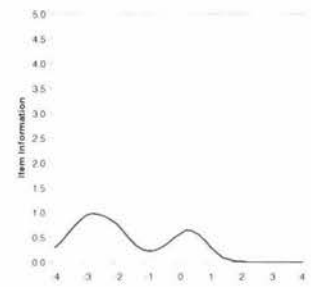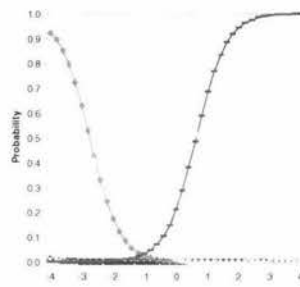
Interpersonal Style   *Tolerance*

Q129   *Treat people as inferiors.*

a	b1	b2	b3	b4
1.95	-3.91	-2.80	-2.02	-0.09
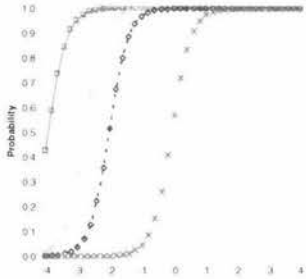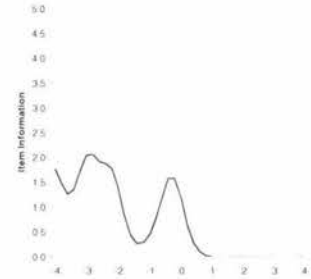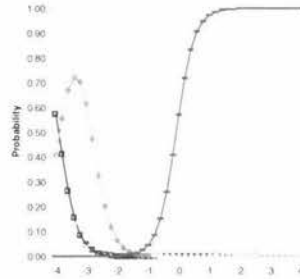
Interpersonal Style   *Tolerance*

Q224   *Respect others.*

a	b1	b2	b3	b4
1.48	-4.91	-4.15	-2.87	-0.39
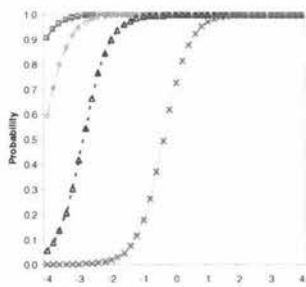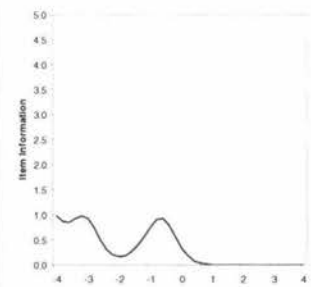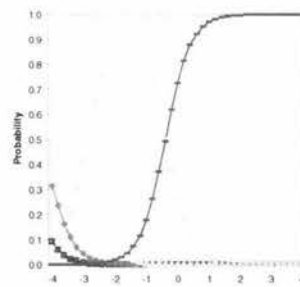
Appendix 5 *(continued)*

Category Characteristic Curve

a = Item discrimination parameter

---□--- b1 = Intersection between option 1 & 2

---◇--- b2 = Intersection between option 2 & 3

···△··· b3 = Intersection between option 3 & 4

×   b4 = Intersection between option 4 & 5

Operating Characteristic Function

---□--- 1- Very Inaccurate

---◇--- 2- Moderately Inaccurate

3- Neither Inaccurate nor Accurate

4- Moderately Accurate

——— 5- Very Accurate

Item Information Function Curve

Interpersonal Style   *Affiliation*

Q075   *Have a sharp tongue.*

a	b1	b2	b3	b4
1.68	-2.83	-1.14	-0.44	0.57
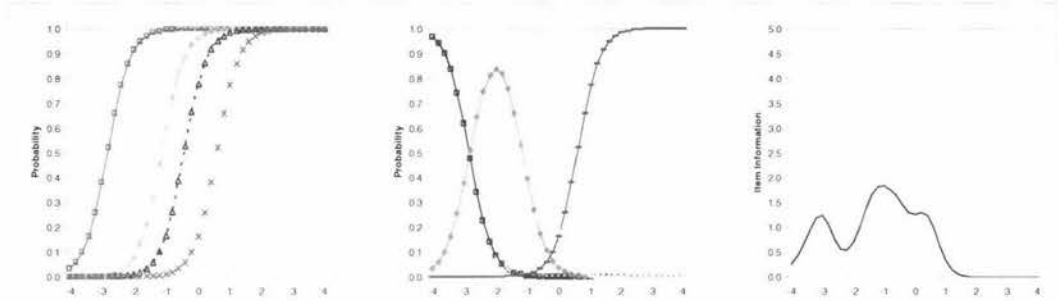
Interpersonal Style   *Affiliation*

Q185   *Insult people.*

a	b1	b2	b3	b4
2.26	-3.45	-2.60	-2.02	-1.06
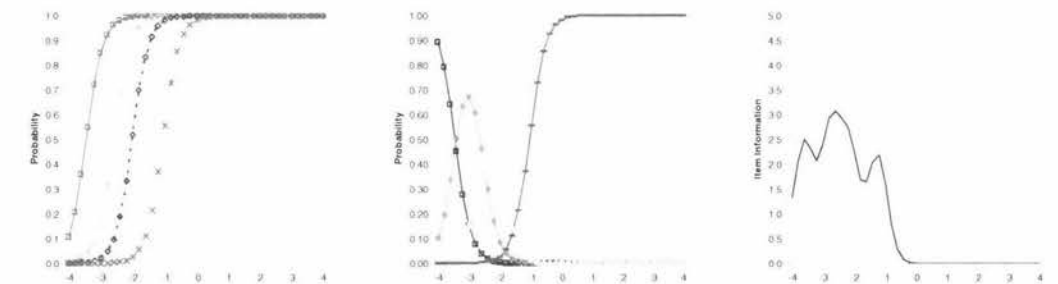
Interpersonal Style   *Affiliation*

Q290   *Yell at people.*

a	b1	b2	b3	b4
1.33	-5.29	-3.37	-2.25	-0.88

Appendix 5 *(continued)*
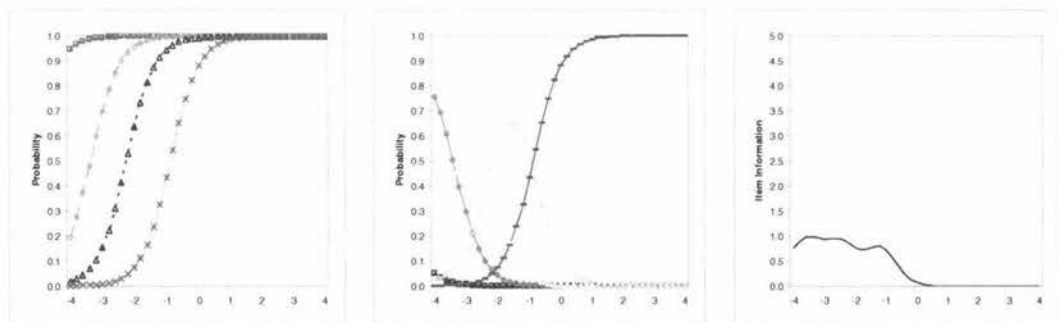
Category Characteristic Curve

a = Item discrimination parameter
⊟ b1 = Intersection between option 1 & 2
b2 = Intersection between option 2 & 3
△ b3 = Intersection between option 3 & 4
× b4 = Intersection between option 4 & 5

Operating Characteristic Function

⊟ 1- Very Inaccurate
◇ 2- Moderately Inaccurate
3- Neither Inaccurate nor Accurate
4- Moderately Accurate
5- Very Accurate

Item Information Function Curve

**Interpersonal Style**   *Empathy*

Q116   *Make people feel welcome.*

a	b1	b2	b3	b4
2.14	-6.26	-3.21	-1.87	0.21
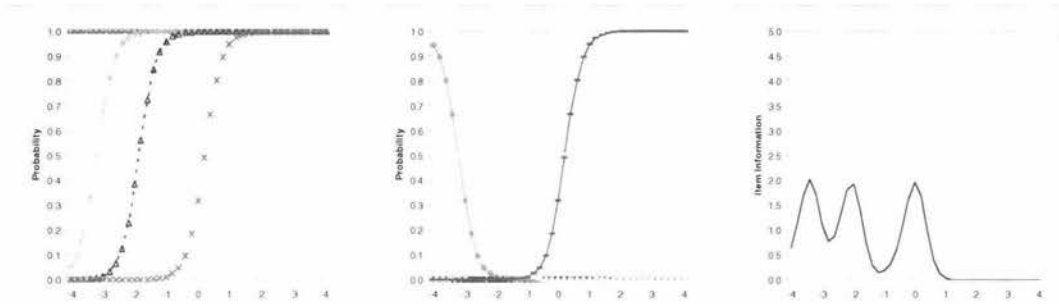
Interpersonal Style   *Empathy*

Q236   *Take time out for others.*

a	b1	b2	b3	b4
2.23	-6.17	-2.54	-1.49	0.48
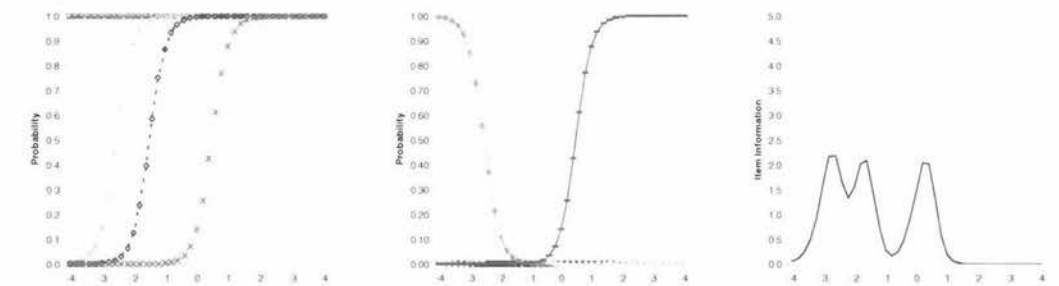
Interpersonal Style   *Empathy*

Q280   *Am concerned about others.*

a	b1	b2	b3	b4
1.71	-7.23	-3.39	-2.12	0.50
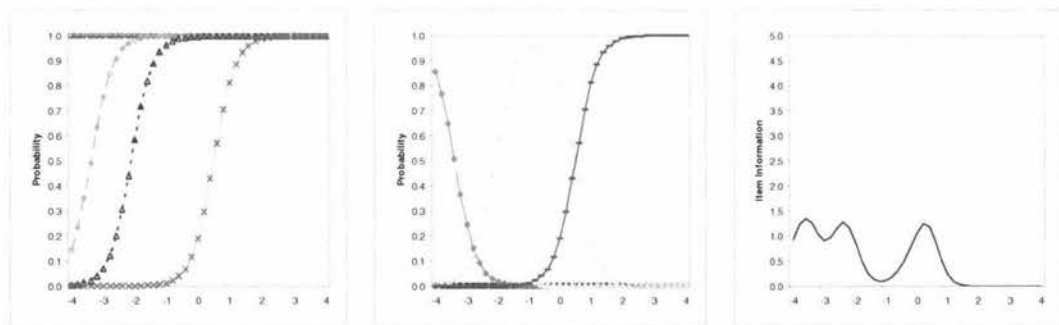
Appendix 5 *(continued)*

Category Characteristic Curve

a = Item discrimination parameter
b1 = Intersection between option 1 & 2
b2 = Intersection between option 2 & 3
b3 = Intersection between option 3 & 4
b4 = Intersection between option 4 & 5

Operating Characteristic Function

1- Very Inaccurate
2- Moderately Inaccurate
3- Neither Inaccurate nor Accurate
4- Moderately Accurate
5- Very Accurate

Item Information Function Curve

Interpersonal Style   *Self-belief*

Q024   *Believe that I am better than others.*

a	b1	b2	b3	b4
1.07	-4.09	-1.1	0.67	1.87
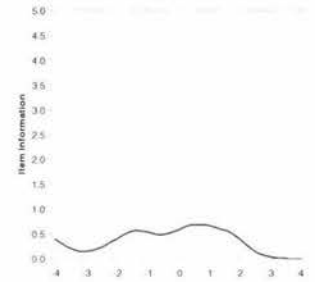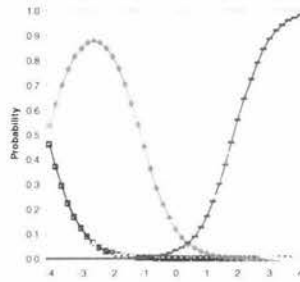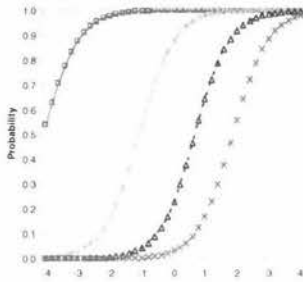
Interpersonal Style   *Self-belief*

Q137   *Have a high opinion of myself.*

a	b1	b2	b3	b4
4.41	-1.39	0.06	0.84	1.77
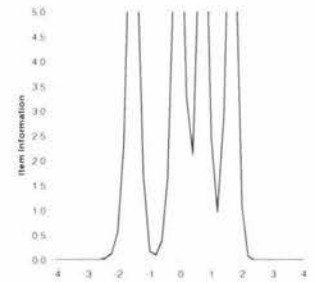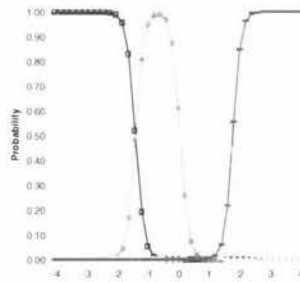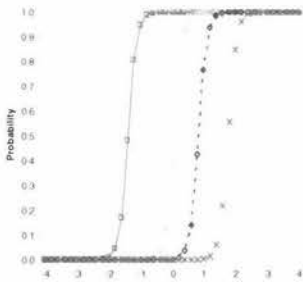
Interpersonal Style   *Self-belief*

Q175   *Think highly of myself.*

a	b1	b2	b3	b4
3.18	-1.03	0.47	1.32	2.12
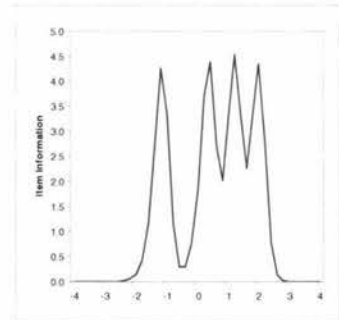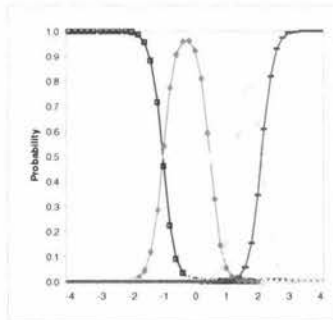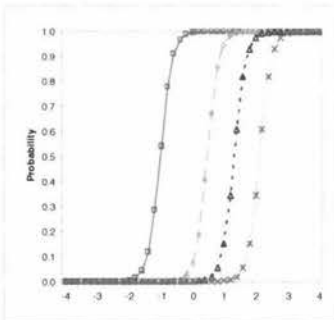
Appendix 5 *(continued)*

Category Characteristic Curve

    a = Item discrimination parameter

—▫— **b1** = Intersection between option 1 & 2

    b2 = Intersection between option 2 & 3

- - -△- - - **b3** = Intersection between option 3 & 4

   × **b4** = Intersection between option 4 & 5

Operating Characteristic Function

—▫— 1- Very Inaccurate

—◇— 2- Moderately Inaccurate

   3- Neither Inaccurate nor Accurate

   4- Moderately Accurate

——— 5- Very Accurate

Item Information Function Curve

Self Management & Drive   *Conscientiousness*

Q033   *Waste my time.*

a	b1	b2	b3	b4
1.61	-3.49	-2.09	-1.16	0.25
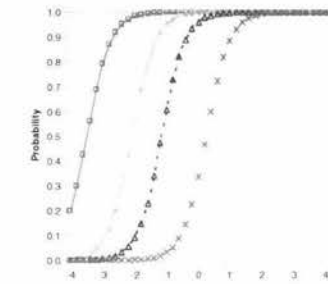
Self Management & Drive   *Conscientiousness*
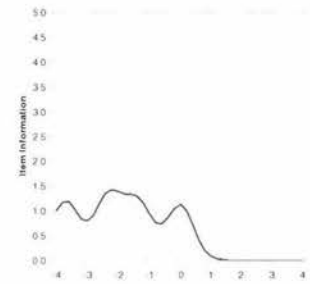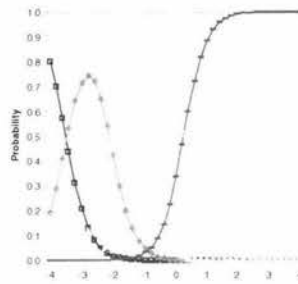
Q233   *Find it difficult to get down to work.*

a	b1	b2	b3	b4
2.00	-3.33	-2.31	-1.53	0.02
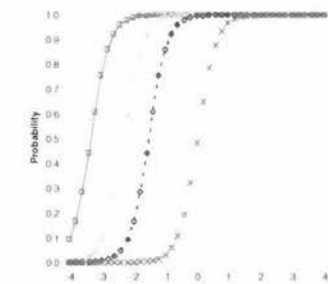
Self Management & Drive   *Conscientiousness*
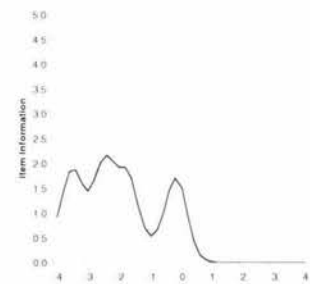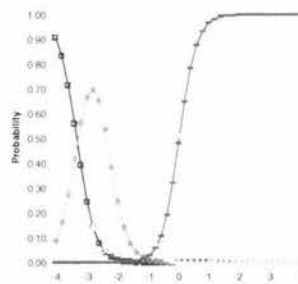
Q259   *Carry out my plans.*

a	b1	b2	b3	b4
2.27	-4.02	-3.40	-2.34	0.16
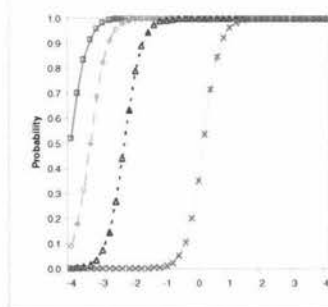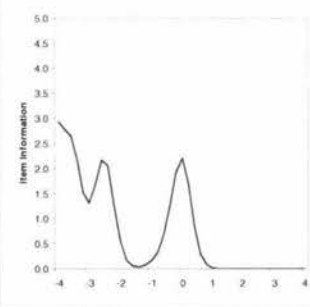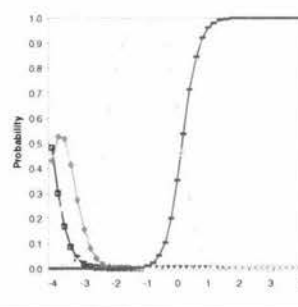
- 92 -

Appendix 5 *(continued)*

Self Management & Drive

Q049 *Am always on the go.*

a	b1	b2	b3	b4
2.17	-3.15	-1.77	-1.05	0.62
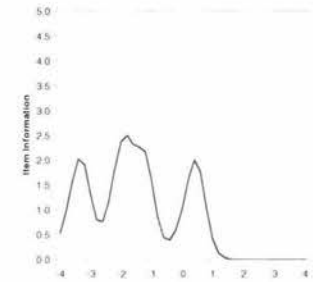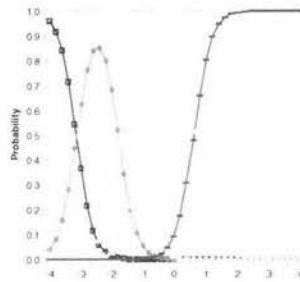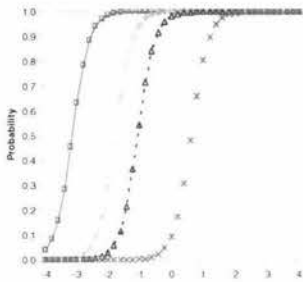
Activity Level

Self Management & Drive

Q127 *Do a lot in my spare time.*

a	b1	b2	b3	b4
1.68	-3.55	-2.07	-1.45	0.24
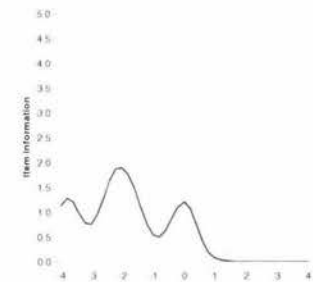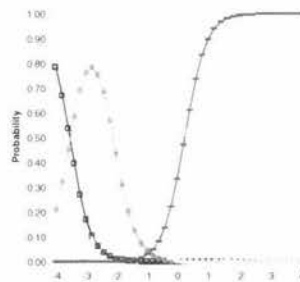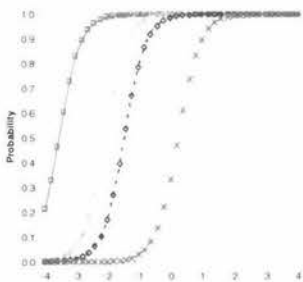
Activity Level

Self Management & Drive

Q130 *Am always busy.*

a	b1	b2	b3	b4
2.46	-3.12	-1.86	-1.23	0.41
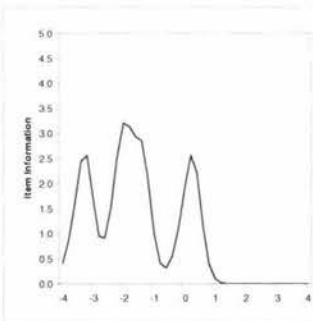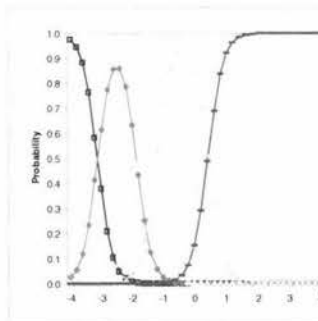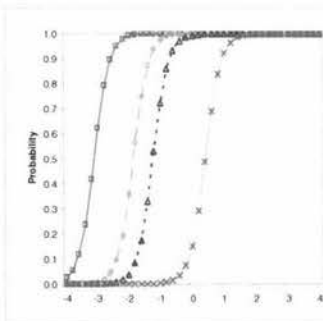
Activity Level

Appendix 5 *(continued)*

Self Management & Drive   *Achievement-Striving*

Q149   *Demand quality.*

a	b1	b2	b3	b4
1.61	-4.65	-3.63	-2.3	0.11

Self Management & Drive   *Achievement-Striving*
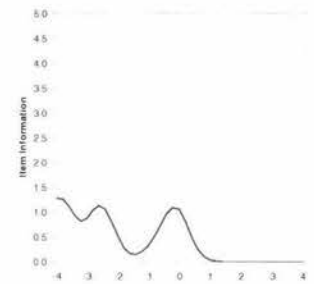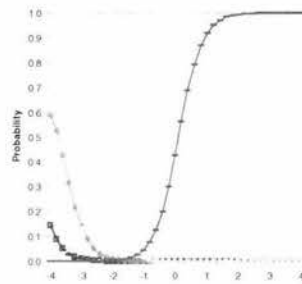
Q229   *Do more than what's expected of me.*

a	b1	b2	b3	b4
1.55	-4.80	-3.54	-1.88	0.38
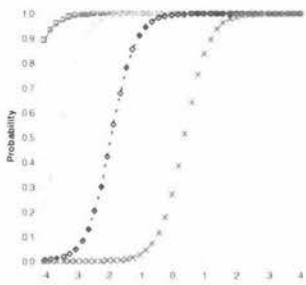
Self Management & Drive   *Achievement-Striving*
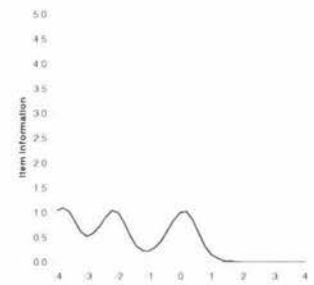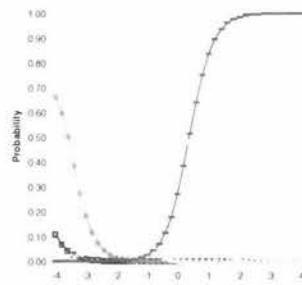
Q315   *Do just enough work to get by.*

a	b1	b2	b3	b4
1.90	-4.02	-2.94	-2.30	0.60
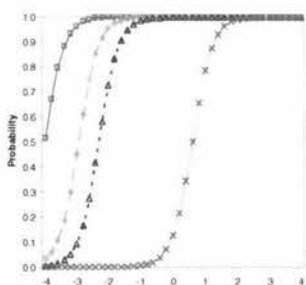
Appendix 5 *(continued)*
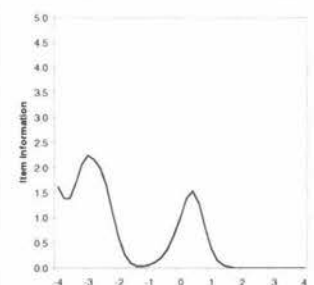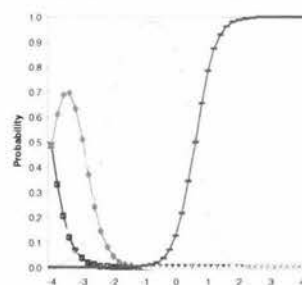
Self Management & Drive   *Self-Discipline*

Q152 *Need a push to get started.*

a	b1	b2	b3	b4
2.86	-2.97	-1.68	-1.11	0.1
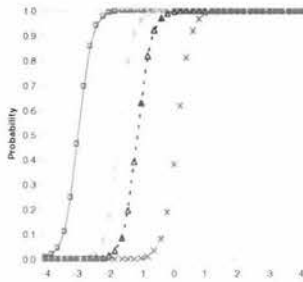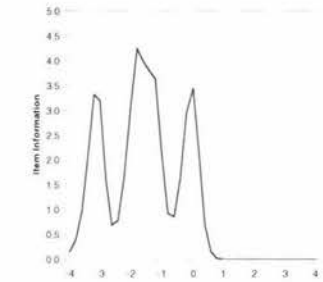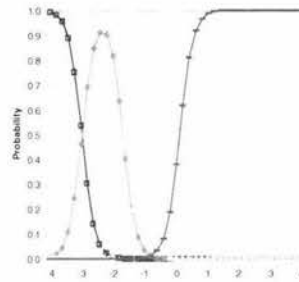
Self Management & Drive   *Self-Discipline*

Q256 *Get to work at once.*

a	b1	b2	b3	b4
1.39	-4.25	-2.25	-1.20	1.05
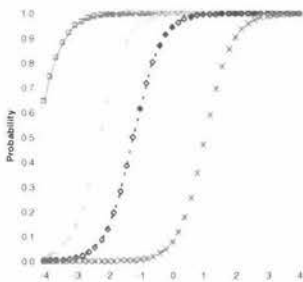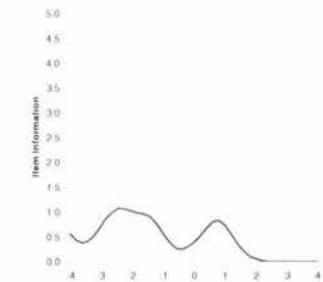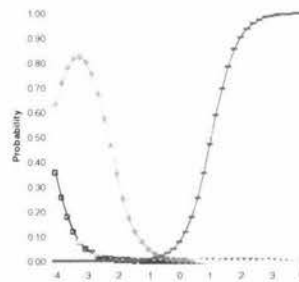
Self Management & Drive   *Self-Discipline*

Q306 *Postpone decisions.*

a	b1	b2	b3	b4
1.60	-3.64	-1.92	-0.99	0.99
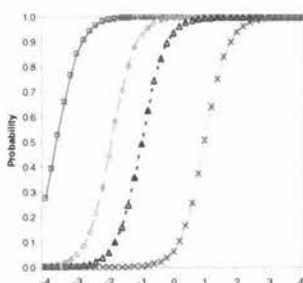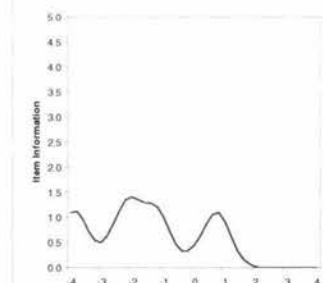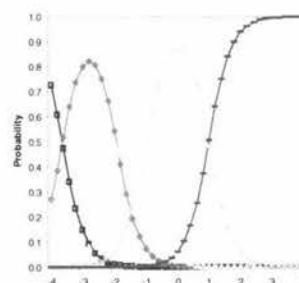
Appendix 5 *(continued)*

Self Management & Drive   *Optimism*

Q003   *Know what I want.*

a	b1	b2	b3	b4
1.2	-5.4	-3.85	-2.39	0.56

Self Management & Drive   *Optimism*

Q197   *Keep improving myself.*

a	b1	b2	b3	b4
2.69	-3.27	-2.46	-1.66	0.21
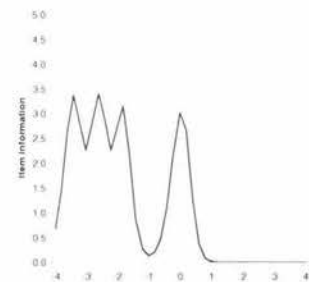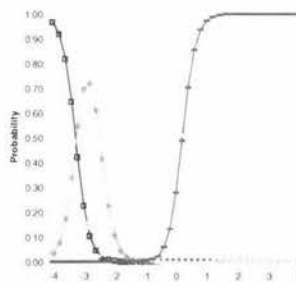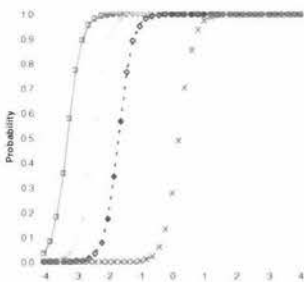
Self Management & Drive   *Optimism*

Q276   *Work on improving myself.*

a	b1	b2	b3	b4
2.95	-5.40	-2.71	-1.94	-0.13
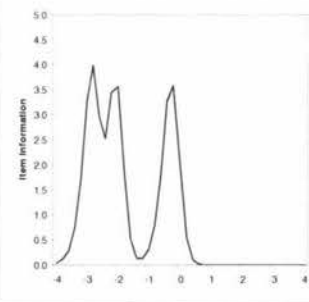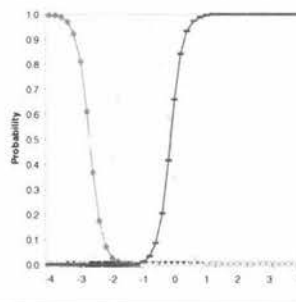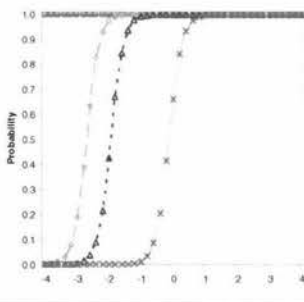
Appendix 5 *(continued)*

Category Characteristic Curve	*Operating Characteristic Function*	*Item Information Function Curve*
a = Item discrimination parameter	1- Very Inaccurate	
b1 = Intersection between option 1 & 2	2- Moderately Inaccurate	
b2 = Intersection between option 2 & 3	3- Neither Inaccurate nor Accurate	
b3 = Intersection between option 3 & 4	4- Moderately Accurate	
b4 = Intersection between option 4 & 5	5- Very Accurate	

Self Management & Drive   *Planned*

Q054   *Like to act on a whim.*

a	b1	b2	b3	b4
1.26	-3.78	-1.47	0.33	1.27
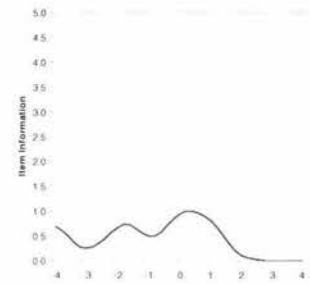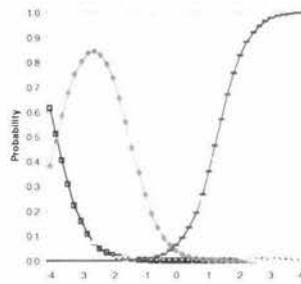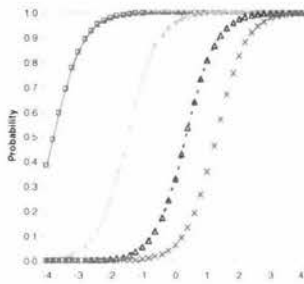
Self Management & Drive   *Planned*

Q195   *Jump into things without thinking.*

a	b1	b2	b3	b4
2.53	-3.35	-1.71	-1.23	0.31
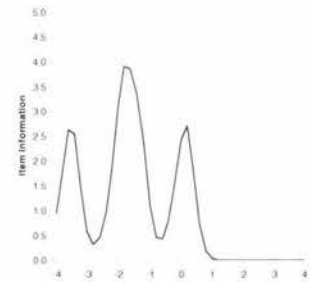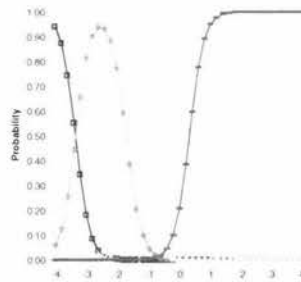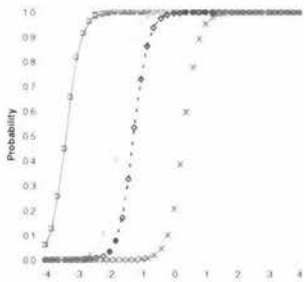
Self Management & Drive   *Planned*

Q264   *Make rash decisions.*

a	b1	b2	b3	b4
2.14	-3.76	-2.30	-1.61	0.09
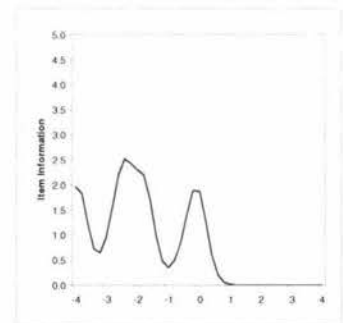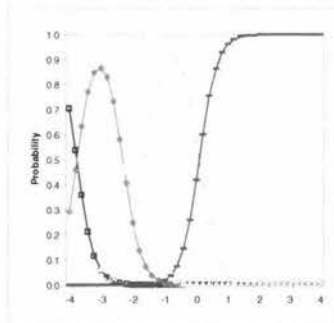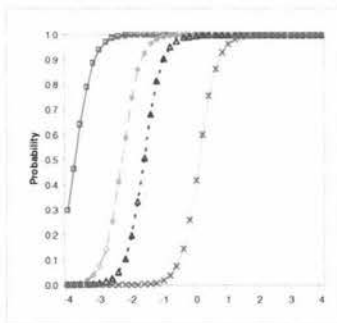
Appendix 5 *(continued)*

Self Management & Drive   *Locus of Control*

Q225   *Habitually blow my chances.*

a	b1	b2	b3	b4
1.44	-5.52	-3.36	-2.36	-0.69
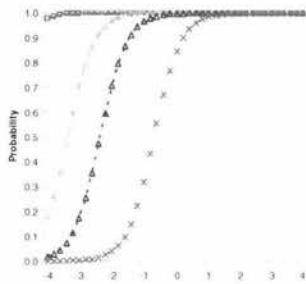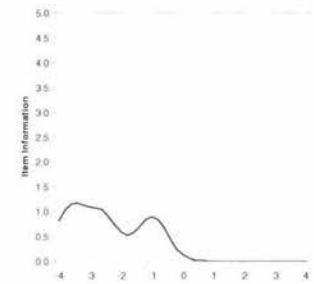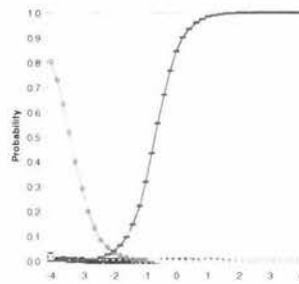
Self Management & Drive   *Locus of Control*

Q244   *Dislike myself.*

a	b1	b2	b3	b4
1.68	-4.14	-3.24	-2.38	-0.97
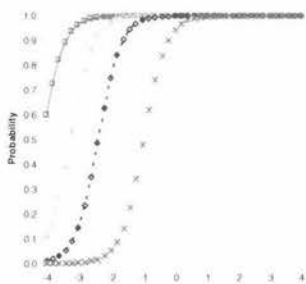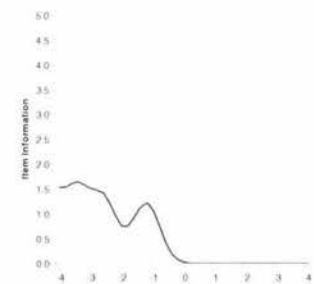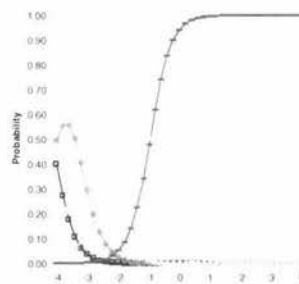
Self Management & Drive   *Locus of Control*

Q249   *Like to take responsibility for making decisions.*

a	b1	b2	b3	b4
1.65	-4.20	-3.47	-2.55	-0.41
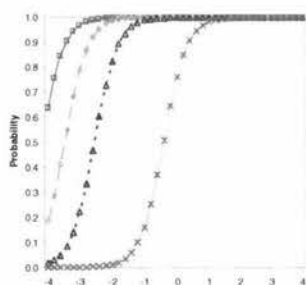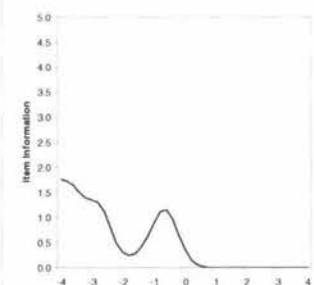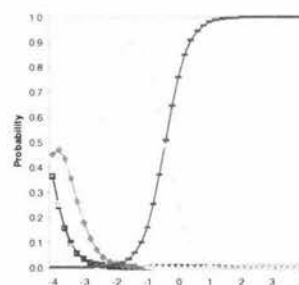
Appendix 5 *(continued)*

Category Characteristic Curve

a = Item discrimination parameter

—☐— b1 = Intersection between option 1 & 2

- - -> - - b2 = Intersection between option 2 & 3

- - -△- - - b3 = Intersection between option 3 & 4

× b4 = Intersection between option 4 & 5

Operating Characteristic Function

—☐— 1- Very Inaccurate

—◇— 2- Moderately Inaccurate

3- Neither Inaccurate nor Accurate

4- Moderately Accurate

—— 5- Very Accurate

Item Information Function Curve

Self Management & Drive   *Compliance*
Q061   *Break rules.*

a	b1	b2	b3	b4
1.77	-3.09	-1	-0.23	1.14
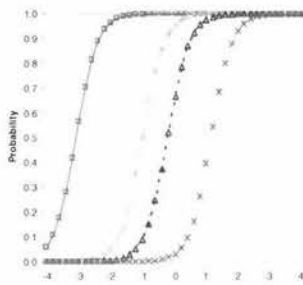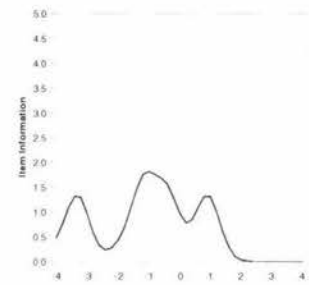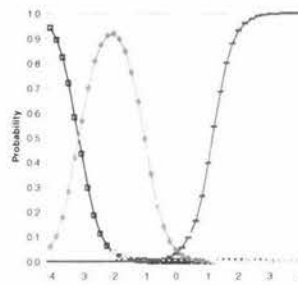
Self Management & Drive   *Compliance*
Q088   *Do the opposite of what is asked.*

a	b1	b2	b3	b4
1.33	-4.96	-4.04	-2.49	-0.87
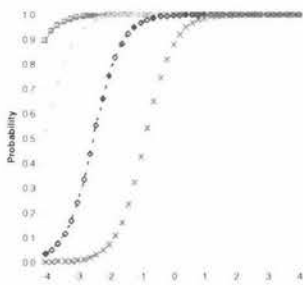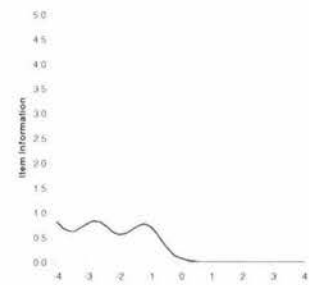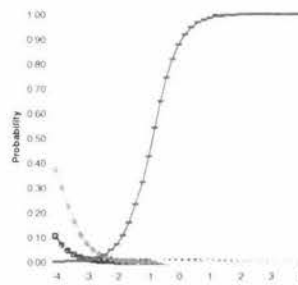
Self Management & Drive   *Compliance*
Q105   *Try to follow the rules.*

a	b1	b2	b3	b4
2.47	-3.17	-1.90	-1.22	0.50
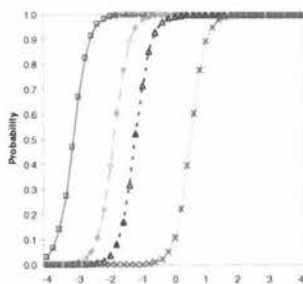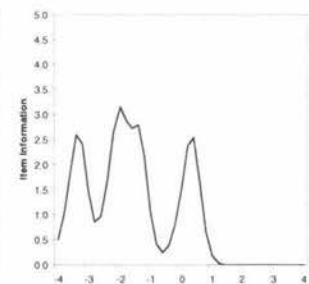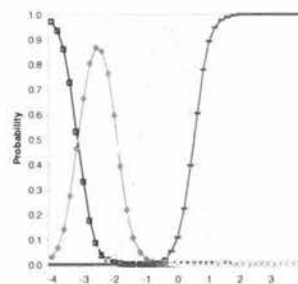
Appendix 5 *(continued)*

Self Management & Drive　　　　*Perfectionism*

Q023　*Want everything to be "just right."*

a	b1	b2	b3	b4
1.79	-3.09	-1.51	-0.62	1.27
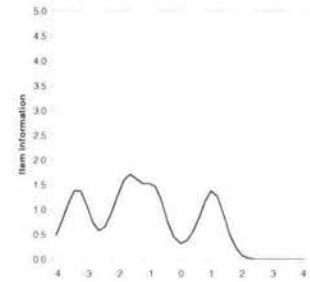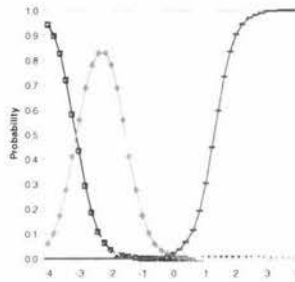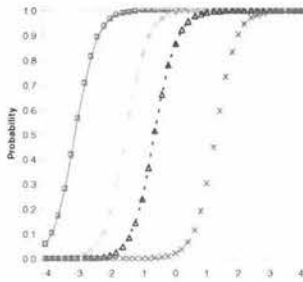
Self Management & Drive　　　　*Perfectionism*

Q269　*Continue until everything is perfect.*

a	b1	b2	b3	b4
2.40	-2.05	-0.60	0.11	2.03
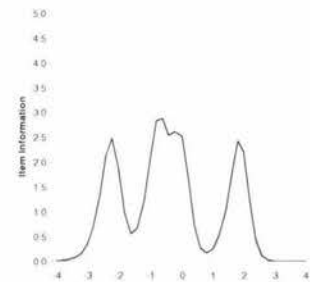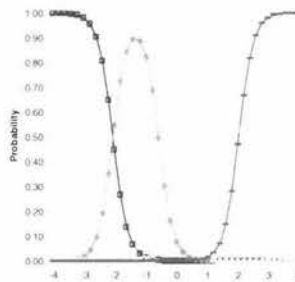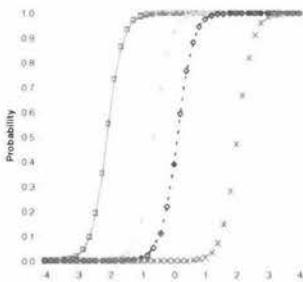
Self Management & Drive　　　　*Perfectionism*

Q271　*Want every detail taken care of.*

a	b1	b2	b3	b4
2.80	-1.94	-0.64	-0.14	1.49

Appendix 5 *(continued)*

Category Characteristic Curve

a = Item discrimination parameter

b1 = Intersection between option 1 & 2

b2 = Intersection between option 2 & 3

b3 = Intersection between option 3 & 4

b4 = Intersection between option 4 & 5

Operating Characteristic Function

1- Very Inaccurate

2- Moderately Inaccurate

3- Neither Inaccurate nor Accurate

4- Moderately Accurate

5- Very Accurate

Item Information Function Curve

Emotional Management   *Anxiety*

Q026   *Get upset easily.*

a	b1	b2	b3	b4
2.01	-0.59	0.97	1.81	3.51

Emotional Management   *Anxiety*

Q032   *Am not easily bothered by things.*

a	b1	b2	b3	b4
1.33	-1.68	0.48	1.38	3.66

Emotional Management   *Anxiety*

Q206   *Get stressed out easily.*

a	b1	b2	b3	b4
2.11	-0.43	1.10	1.80	3.14
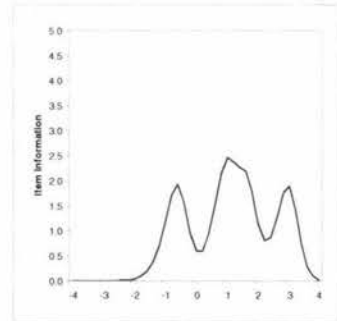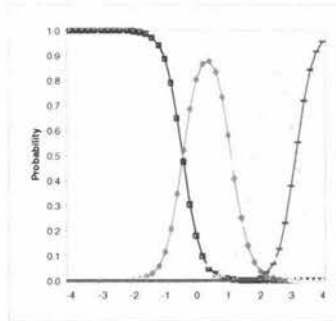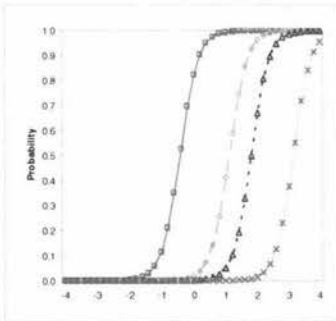
Emotional Management   *Approval Seeking*

Q016   *Worry about what people think of me.*

a	b1	b2	b3	b4
1.97	-2.46	-1.09	-0.28	1.64
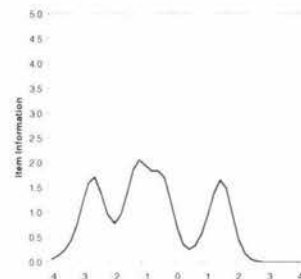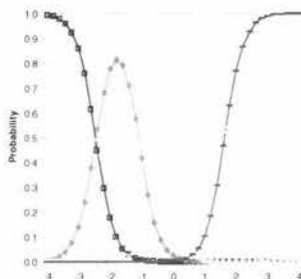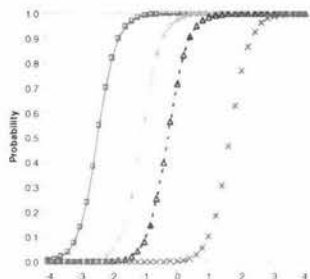
Emotional Management   *Approval Seeking*

Q062   *Need the approval of others.*

a	b1	b2	b3	b4
1.86	-1.81	-0.45	0.34	2.43
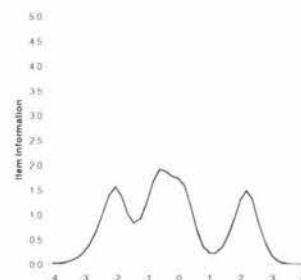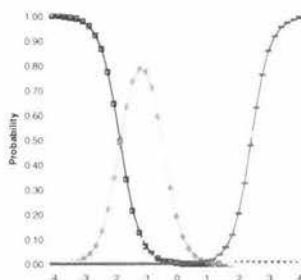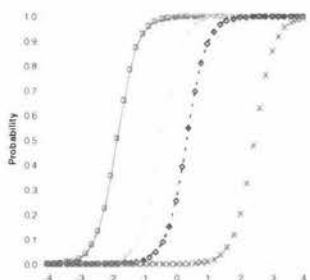
Emotional Management   *Approval Seeking*

Q222   *Want to amount to something special in others' eyes.*

a	b1	b2	b3	b4
1.13	-3.52	-1.85	-0.55	1.61
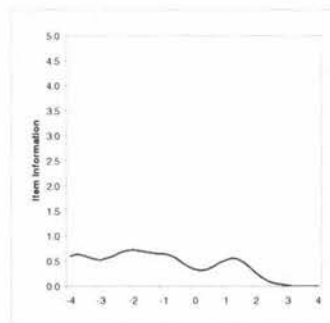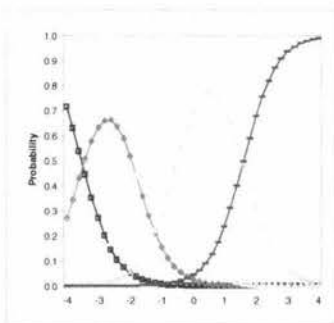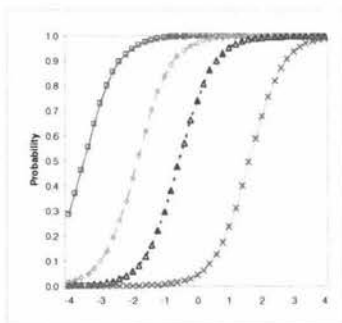
Appendix 5 *(continued)*

Category Characteristic Curve

    a = Item discrimination parameter

⎯□⎯ **b1** = Intersection between option 1 & 2

⎯◇⎯ **b2** = Intersection between option 2 & 3

- - △- - - **b3** = Intersection between option 3 & 4

  ✕   **b4** = Intersection between option 4 & 5

Operating Characteristic Function

⎯□⎯ 1- Very Inaccurate

⎯◇⎯ 2- Moderately Inaccurate

3- Neither Inaccurate nor Accurate

4- Moderately Accurate

⎯ 5- Very Accurate

Item Information Function Curve

Emotional Management   *Emotional Responsiveness*

Q131   *Think about the causes of my emotions.*

a	b1	b2	b3	b4
2.3	-2.69	-1.42	-0.85	0.84
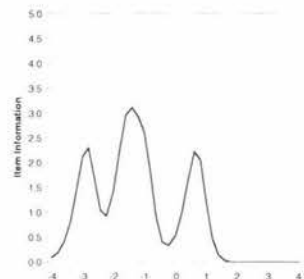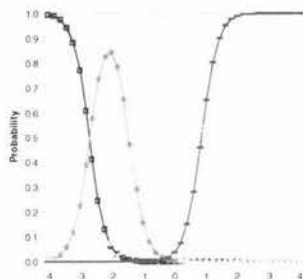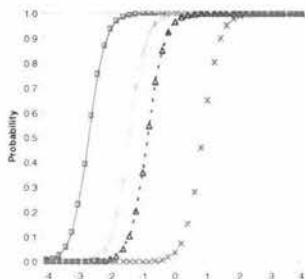
Emotional Management   *Emotional Responsiveness*

Q144   *Pay a lot of attention to my feelings.*

a	b1	b2	b3	b4
2.17	-2.79	-1.12	-0.07	1.62
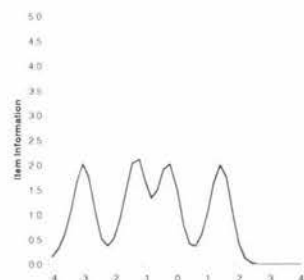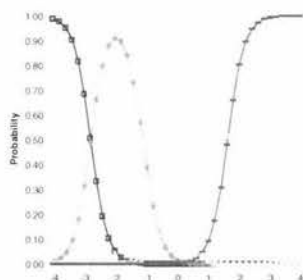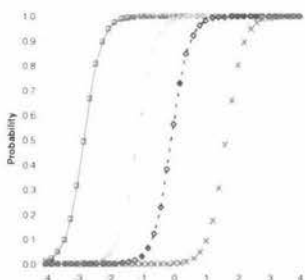
Emotional Management   *Emotional Responsiveness*

Q171   *Notice my emotions.*

a	b1	b2	b3	b4
3.24	-3.19	-1.69	-0.95	0.79
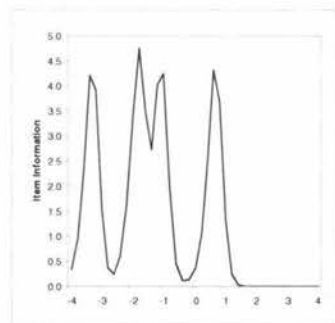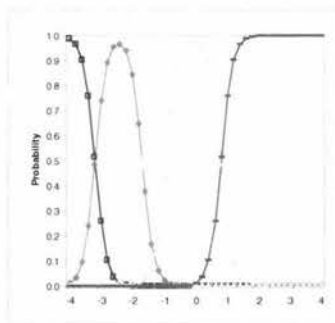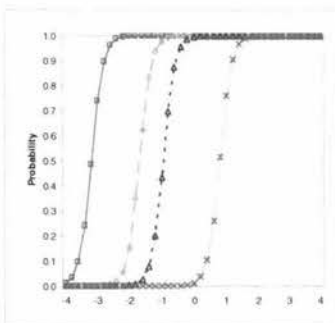
Appendix 5 *(continued)*

Emotional Management   *Emotional Expression*

Q190   *Am not easily affected by my emotions.*

a	b1	b2	b3	b4
1.91	-1.86	0.03	0.74	2.68
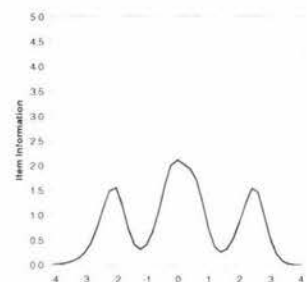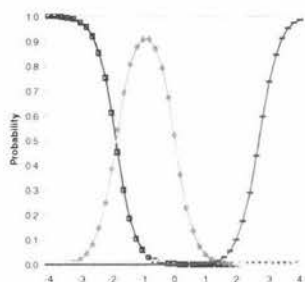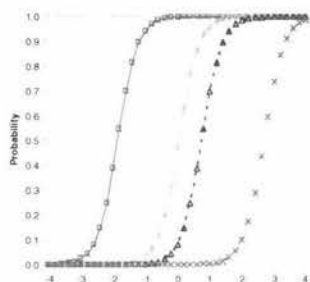
Emotional Management   *Emotional Expression*

Q246   *Experience my emotions intensely.*

a	b1	b2	b3	b4
1.35	-1.59	0.02	1.09	3.03
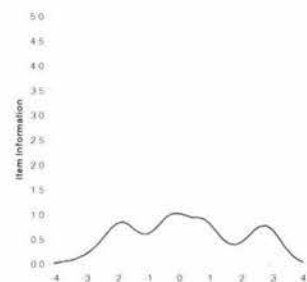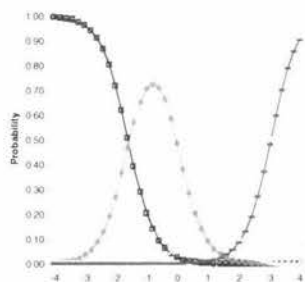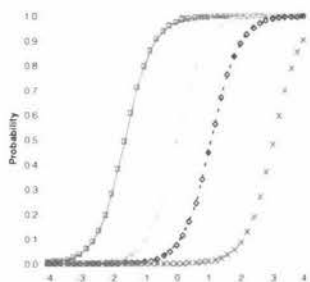
Emotional Management   *Emotional Expression*

Q258   *Seldom get emotional.*

a	b1	b2	b3	b4
1.93	-1.37	0.28	0.89	2.54
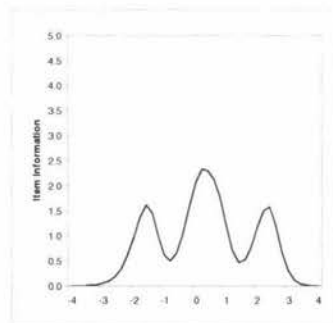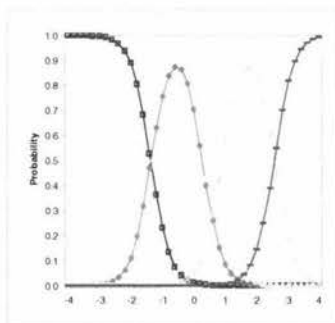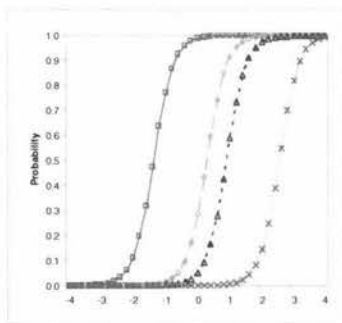
Appendix 6

Full Account of the Development Process for the Lie Scale, Desirable Virtues

The original state of the lie scale is shown below in Table 1.

Table 1

Original State of the Lie Scale Desirable Virtues

Scale	Item	Pos or Neg	Wording
Desirable Virtues			
	Q037	+	Always admit it when I make a mistake.
	Q077	+	Never give up hope.
	Q148	+	Know that anyone who tries can get a job.
	Q285	+	Always know why I do things.
	Q213	+	Never give up.
	Q219	+	Know immediately what to do.
	Q239	+	Believe there is never an excuse for lying.
	Q108	+	Always know what I am doing.
	Q042	+	Am always ready to start afresh.
	Q092	+	Have never engaged in gossip.
	Q293	+	Will do anything for others.
	Q147	+	Am always prepared.
	Q034	−	Don't always practice what I preach.
	Q084	−	Have some bad habits.
	Q098	−	Have sometimes had to tell a lie.
	Q313	−	Am not always honest with myself.
	Q076	−	Am not always what I appear to be.

Of these 17 items, 6 items were deleted as part of the unidimensionality analysis (see Table 2).

Table 2

Lie scale item deletions for unidimensionality.

Scale	Deletions	Item	Squared Multiple Correlation
Desirable Virtues	6	Q148	0.014
		Q293	0.045
		Q037	0.064
		Q042	0.066
		Q077	0.072
		Q076	0.085

The fit statistics and reliability alpha for this scale are shown in Table 3.

Table 3

Fit Statistics and Reliability Alpha for the Lie Scale

Scale	Items	Alpha	χ^2	DF	Prob	GFI	TLI	CFI	RMSEA
Desirable Virtues	11	0.72	215.00	44	0.00	0.92	0.70	0.76	0.09

MultiLog was then used to extract the 'a' and 'b' parameters for this scale (Table 4).

Table 4

Parameter Estimates for the Lie Scale

Scale	Item	a	b1	b2	b3	b4	Wording
Desirable Virtues							
	Q034	1.22	-3.96	-1.35	-0.60	1.21	Don't always practice what I preach.
	Q084	0.79	-2.59	0.72	2.04	4.24	Have some bad habits.
	Q092	0.70	-1.82	1.23	2.42	5.82	Have never engaged in gossip.
	Q098	0.70	-2.61	0.58	1.45	3.26	Have sometimes had to tell a lie.
	Q108	1.47	-3.11	-1.64	-0.78	1.76	Always know what I am doing.
	Q147	1.35	-4.59	-2.10	-1.04	1.89	Am always prepared.
	Q213	0.96	-6.18	-3.95	-2.63	0.29	Never give up.
	Q219	1.04	-5.24	-2.04	-0.27	2.81	Know immediately what to do.
	Q239	0.84	-3.57	-0.90	-0.05	2.32	Believe there is never an excuse for lying.
	Q285	1.42	-4.62	-2.60	-1.51	0.92	Always know why I do things.
	Q313	1.02	-4.57	-2.00	-1.09	0.88	Am not always honest with myself.

Using Samejima's GR model (1969) and *Microsoft Excel*, each of these items were analysed according to the properties mentioned previously in this study. Table 5 shows the items that were selected along with the new reliability alpha for this scale.

Table 5

Parameter Estimates for the Three Items Selected to for the Lie Scale

Scale	Item	a	b1	b2	b3	b4	Wording
Desirable Virtues							
	Q108	1.47	-3.11	-1.64	-0.78	1.76	Always know what I am doing.
	Q147	1.35	-4.59	-2.10	-1.04	1.89	Am always prepared.
	Q285	1.42	-4.62	-2.60	-1.51	0.92	Always know why I do things.
Alpha:	0.611						

Fig. 1 illustrates a graphical representation of these three items.

Fig. 1

CCC's, OCF's, and IIF's for the Three Lie Scale Items Selected for Desirable Virtues

Category Characteristic Curve	Operating Characteristic Function	Item Information Function Curve
a = Item discrimination parameter	1- Very Inaccurate	
b1 = Intersection between option 1 & 2	2- Moderately Inaccurate	
b2 = Intersection between option 2 & 3	3- Neither Inaccurate nor Accurate	
b3 = Intersection between option 3 & 4	4- Moderately Accurate	
b4 = Intersection between option 4 & 5	5- Very Accurate	

Desirable Virtues   *Desirable Virtues*

Q108   *Always know what I am doing.*

a	b1	b2	b3	b4
1.47	-3.11	-1.64	-0.78	1.76
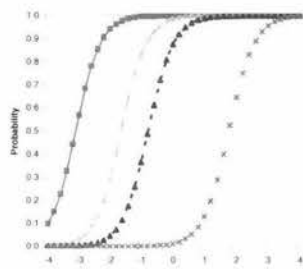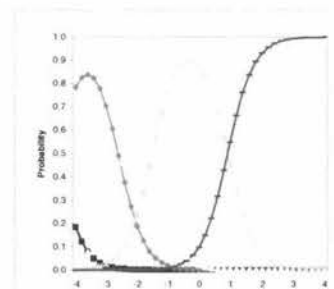
Desirable Virtues   *Desirable Virtues*

Q147   *Am always prepared.*

a	b1	b2	b3	b4
1.35	-4.59	-2.10	-1.04	1.89
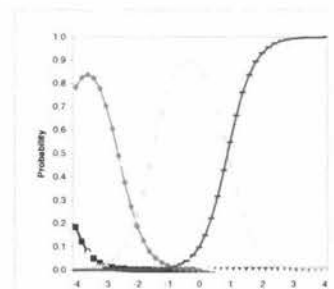
Desirable Virtues   *Desirable Virtues*

Q285   *Always know why I do things.*

a	b1	b2	b3	b4
1.42	-4.62	-2.60	-1.51	0.92