The Spanish Journal of Psychology (2018), **21**, e63, 1–13.
© Universidad Complutense de Madrid and Colegio Oficial de Psicólogos de Madrid
doi:10.1017/sjp.2018.62

CAMBRIDGE UNIVERSITY PRESS

Validation of a Questionnaire for Personality Profiling using Cognitive Diagnostic Modeling

Javier Revuelta[1], Lucia Halty[2] and Carmen Ximénez[1]

[1] *Universidad Autónoma de Madrid (Spain)*
[2] *Universidad Pontificia Comillas (Spain)*

Abstract. This article describes the development of the ENCUIST (Extroversion, Neuroticism, Callous-Unemotional, Instability, Short-Test) questionnaire, which has been created to provide a personality profiling method based on a cognitive diagnostic modeling framework. The ENCUIST measures the attributes of extroversion, neuroticism, callous unemotionality and overt expressions of anger that are relevant in a forensic context. The scores provided by the ENCUIST are binary classifications of the individuals (high/low) in these attributes. The ENCUIST was developed using a sample of 516 subjects to study its validation through psychometric procedures, including factor analysis, cognitive diagnostic modeling and structural equation modeling. The results supported a four-factor structure. Linear regressions were used to evaluate the predictive validity of the scores provided by ENCUIST with respect to two external criteria that are relevant in the forensic context, namely behavioral activation and behavioral inhibition. The results showed that the extroversion dimension is positively related to behavioral activation, although the effect size is modest and the proportion of explained variance is only 11%. Moreover, the dimensions of neuroticism and anger expression are positively related to behavioral inhibition, with 7% of the variance explained. Together, these results suggest that cognitive diagnostic models are useful tools for the elaboration of personality profiles based on classifying subjects along binary attributes.

Received 17 August 2017; Revised 4 November 2018; Accepted 6 November 2018

Keywords: Cognitive Diagnostic Models, factor analysis, forensic profiling, personality profile, structural equation models.

One important application field for personality theory is forensics, in which the purpose is to detect individuals who score high or low in certain personality traits that are relevant for anti-social behavior. Personality profiling of witnesses and defendants is a common application of psychology in criminal cases. The application of personality profiling questionnaires is part of a larger evaluation process that is based mainly on interviews. For that reason, questionnaires cannot be based on large numbers of items, as is typical in personality testing. Rather, questionnaires need to be developed that can provide an individual classification on relevant psychological attributes using a limited number of items.

The present article describes the development of a questionnaire, the ENCUIST, that can be useful in a forensic context. The ENCUIST provides a binary classification of individuals on several personality traits, which determine the measurement model used to validate the questionnaire. The process of elaborating the ENCUIST consists of selecting a few items with good psychometric properties from a larger set of items to which a calibrating sample responded.

The current state of the art in the field of personality assessment maintains that personality traits are continuous variables (Haslam, Holland, & Kuppens, 2012). However, we are developing a questionnaire based on binary classification because it is useful in the particular cases of the forensic applications. Information about personality characteristics of persons participating in forensic process is usually obtained from interviews. ENCUIST will be used in this context for the interviewers to identify relevant pieces of information that help to classify individuals as high or low in the personality attributes. In the forensic context there is a need to obtain a personality profile to orient further investigations. This profile has to be obtained quickly using a check-list and the evaluation of personality by applying a long test and dichotomization of scale scores is too time-consuming.

Correspondence concerning this article should be addressed to Javier Revuelta. Universidad Autónoma de Madrid. Facultad de Psicología. Departamento de Psicología Social y Metodología. 28049 Madrid (Spain).

E-mail: javier.revuelta@uam.es

This work was supported by Grant PSI2015-66366-P from the Spanish Ministerio de Ecomomía y Competitividad. The data were analyzed using the resources of the Centro de Computación Científica of the Universidad Autónoma de Madrid (CCC, https://www.ccc.uam.es).

How to cite this article:
Revuelta, J., Halty, L., & Ximénez, C. (2018). Validation of a Questionnaire for Personality Profiling using Cognitive Diagnostic Modeling. *The Spanish Journal of Psychology, 21*. e63. Doi:10.1017/sjp.2018.62

https://doi.org/10.1017/sjp.2018.62 Published online by Cambridge University Press

The present investigation is part of a larger project. The work described here consists of the calibration of ENCUIST using data from the general population. We did so because the traits measured by the ENCUIST are present in both general and forensic populations. Thus the data from the general population provides relevant information about the structure of the questionnaire. Notwithstanding, this investigation shall be followed by a detailed investigation of the validity of the ENCUIST scores using smaller samples in the forensic context.

The measurement framework to validate the ENCUIST makes use of Cognitive Diagnostic Models (CDMs). CDMs are a family of measurement models that classify individuals with respect to categorical latent attributes (de la Torre, van der Ark, & Rossi, 2015; Sorrel et al., 2016). The attributes usually are dichotomous, and the dichotomies can be interpreted as the presence or absence of the attribute. CDMs originated in the field of educational measurement and have been extended to the analysis of situational judgments (Garcia, Olea, & de la Torre, 2014; Sorrel et al., 2016) and clinical data (de la Torre et al., 2015). In the present article we extend this approach to the development of personality profiles, represented by the patters of binary attributes.

The article is organized as follows. First, substantive aspects are discussed, including explanations of the personality traits relevant for forensic contexts and the selection of items measuring these traits from which the ENCUIST is constructed. Second, we present a brief introduction to CDMs because of the novelty of this approach to the field of personality. Third, the results of three data analyses are presented. The first analysis describes the application of factor analysis to investigate the dimensions measured by the ENCUIST. The second analysis is the application of CDM technique to estimate the personality profiles and the results of a linear regression analysis to predict exogenous variables from the profiles. The third analysis is the application of a structural equation model to compare its results with those of the CDM and to illustrate the similarities and differences between the two approaches. The article concludes with a general discussion summarizing the results and their practical implications.

Personality traits measured by the ENCUIST questionnaire

A reference personality model is needed to select the attributes that constitute the profile. The ENCUIST is based on Eysenck's (1952) model of personality because it is a well-known consolidated taxonomy in this field. However, some modifications will be made to include some traits relevant for forensic contexts. Eysenck's model defines three essential personality traits: *Psychoticism* (*P*), *Extraversion* (*E*), and *Neuroticism* (*N*). Eysenck established a hierarchical personality structure in which the three traits (*PEN*) are located at the top, and within each one are more specific factors corresponding to people's common behavioral responses. Of the three traits of the model, E and N have received the most support and have even been retained in other personality models, such as the Big Five (Goldberg, 1993), where the P trait has been least replicated in subsequent investigations (Corr, 2010). For these reasons, we replaced the P trait with the *Callous Unemotional* trait or *CU*, which is a trait that Eysenck wished to reflect in his P dimension and is potentially more relevant in forensic contexts (Frick & White, 2008). Subjects with high scores on CU have difficulties recognizing and reacting to certain emotions in others, specifically fear (Blair et al., 2004), and they have serious difficulties processing facial expressions of fear or sadness. That is, they have longer reaction times for recognizing such emotional expressions, and when they say they recognize them, they commit more identification mistakes than subjects without these traits (Blair, Colledge, Murray, & Mitchell, 2001). Good processing of facial expressions is crucial for socialization and modulation of interpersonal behavior. Knowing how to identify others' emotions allows one to act accordingly and curb behaviors that can cause interpersonal damage.

In addition to CU, we considered *Expression of Anger* (*EA*) as another important trait to take into account. Anger is a universal emotion that everyone experiences at some time, but some people feel it with excessive frequency and intensity and, for this reason it may be dysfunctional (Jasisnki et al., 2016). High levels of anger are associated with aggressive behaviors (Webster et al., 2015), domestic violence (Barbour, Eckhardt, Davison, & Kassinove, 1998), health problems (Smith, Glazer, Ruiz, & Gallo, 2004), and interpersonal problems (Baron et al., 2007). Given the close relation between anger and antisocial behavior, we think it is appropriate to introduce it into the theoretical model and relate it to *Emotional Instability*. That is, a person will be considered emotionally unstable if he/she has high levels of fear/anxiety (Eysenck's original conception), high levels of anger, or the presence of both.

In summary, the traits that will be taken into account in the theoretical model to establish personality profiles are *Extroversion* (E), *Neuroticism* (N), *Callous Unemotional* (CU), and *Expression of Anger* (EA). We elaborated the ENCUIST to assess these four traits. The selected traits have been shown to be related to several personality traits in different models. We will relate the variables of the ENCUIST with the variables of Gray's (1970) model of the biological bases of personality traits, which is the best-known alternative to the Eysenck model (Matthews & Gilliland, 1999). Gray's model

https://doi.org/10.1017/sjp.2018.62 Published online by Cambridge University Press

distinguishes two types of psychobiological systems: The first, the *Behavioral Activation System* (*BAS*), responds to appetitive conditioned signals (reward and non-punishment), activates approach behavior towards these types of signals or stimuli, and guides the organism towards the goals it needs to obtain. The BAS has been associated with positive moods, and its over-activation is involved in impulsivity (Meyer & Hofmann, 2005; Newman, MacCoon, Vaughn, & Sadeh 2005). Regarding the model we will use herein, we expect to find a positive relation between the BAS and E. The second system, the *Behavioral Inhibition System* (*BIS*), is a system that responds to aversive conditioned stimuli, that is, to signs of punishment and non-reward, or to unfamiliar stimuli, by suppressing or inhibiting behavior. The BIS has been linked to the development of negative moods and anxiety (Newman et al., 2005) and, for this reason, we expect that high scores on the BIS will be related to EA. On another hand, hypoactivity of the BIS (low BIS) is related to CU because, as mentioned, people with high scores on CU have difficulties experiencing emotions related to fear and anxiety (Heym & Lawrence, 2010).

Cognitive Diagnostic Models

Cognitive diagnostic modeling is used to study the relation between underlying categorical psychological attributes and responses to questionnaire items (de la Torre, 2009). The vast majority of CDM models assume binary attributes, although there are also some examples of polytomous attributes (Chen & de la Torre, 2013). This paper considers only binary attributes because they are more relevant to the forensic context in which ENCUIST will be applied. The binary classification of the subject provided by the CDM is interpreted as being high or low on the personality variables.

From a statistical viewpoint, CDMs are structured latent class models that classify subjects into latent categories based on their item response patterns (Maris, 1999). The key difference between CDMs and other popular psychometric models, such as factor analysis and item response theory, is that individuals are represented by categorical attributes in CDMs and by continuous traits in the other models.

A large number of CDMs have been described in the psychometric literature, among these we have selected the G-DINA model (*Generalized Deterministic Input, Noisy and Gate*) because it is a versatile model, allowing the assessment of the effect of each attribute on each item, the interactions among the attributes within the item, and the interactions between attributes in the general population of individuals (de la Torre, 2011). The G-DINA model applies to dichotomous observed responses, where the response categories are labeled *agree/disagree*, *right/wrong*, etc.

The input to G-DINA is a matrix of responses and the so-called Q-matrix that represent the relation between observed items and latent attributes (Embretson, 1984; Tatsuoka, 1985). The Q-matrix has order $J \times K$, where J is the number of items and K is the number of attributes. The entries of the Q-matrix are zeros and ones. The intersection between row j and column k has the value of 1 when the attribute k is measured by the item j, and of 0 otherwise.

G-DINA represents the individual i by a profile, $\boldsymbol{\alpha}_i = (\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{iK})$. The elements of the profile are binary values that indicate whether the subject scores high or low on each of the attributes measured by the test. In general, the number of distinct profiles for a test that measures K attributes is 2^K. For example, the number of profiles for a test that measures four attributes is 2^4=16. The profile $\boldsymbol{\alpha}_i = (0000)$ indicates that the subject scores low on all four attributes, the profile $\boldsymbol{\alpha}_i = (1110)$ indicates that the subject scores high on the first three attributes and low on the fourth, etc. The $\boldsymbol{\alpha}_i$ profile is not directly observable and, for this reason, it is estimated from the responses to the questionnaire.

We shall examine the notation of the G-DINA model by means of the specific model that will be applied later in our empirical example. Suppose that observed responses are coded with 0 and 1, X_{ij} denotes the response of subject i to item j and the item measures two attributes. The log-odds of the probability of a positive response is given by

$$\log \frac{P(X_{ij}=1 \mid \boldsymbol{\alpha}_i)}{P(X_{ij}=0 \mid \boldsymbol{\alpha}_i)} = \delta_{j0} + \delta_{j1}\alpha_{i1} + \delta_{j2}\alpha_{i2} + \delta_{j12}\alpha_{i1}\alpha_{i2}, \quad (1)$$

where the parameter δ_{j0} is the baseline and represents the log-odds of the positive response when the subject possesses neither the attributes α_1 nor α_2. Parameters δ_{j1} and δ_{j2} represent the increase in the log-odds as a result of the presence of the attribute α_1 or α_2; and parameter δ_{j12} represents the interaction of attributes within the item. This model is referred to as *G-DINA with a logit link* in the psychometric literature.

The second part of the G-DINA model is the probability of the profiles. Those profiles follow a categorical distribution, according to which the probability of the profile $\boldsymbol{\alpha}_r$ (where $r = 1, \ldots, 2^K$) is

$$P(\boldsymbol{\alpha}_r) = \pi_r. \quad (2)$$

The parameters of the model in Equation (2) are the probabilities of 2^K profiles: $\pi_1, \ldots, \pi_r, \ldots, \pi_{2^K}$. Main effects and interactions between attributes are studied by fitting a log-linear model (with logit link) to the probabilities in Equation (2). More specifically, for a test that

https://doi.org/10.1017/sjp.2018.62 Published online by Cambridge University Press

measures four attributes the probabilities in Equation (2) are computed as

$$\begin{aligned}\text{logit}\pi_r = \omega_0 + \sum_{k=1}^{K}\alpha_{rk}\omega_k + \sum_{k=1}^{K}\sum_{l=1}^{K}\alpha_{rk}\alpha_{rl}\omega_{kl} \\ + \sum_{k=1}^{K}\sum_{l=1}^{K}\sum_{m=1}^{K}\alpha_{rk}\alpha_{rl}\alpha_{rm}\omega_{klm} + \alpha_{rk}\alpha_{rl}\alpha_{rm}\alpha_{rn}\omega_{klmn}\end{aligned} \quad (3)$$

The parameters determine the pattern of associations among the attributes; ω_0 is the intercept that represents the value of the logit when all the attributes take zero value, ω_k is the main effect of the attribute α_k, ω_{kl} is the interaction parameter between a pair of attributes, ω_{klm} is the interaction among three attributes, and ω_{klmn} is the interaction between the four attributes. In the more general case, when there are interactions of any order between attributes, model (3) has $2^k - 1$ parameters. Particular cases of Equation (3) can be applied to investigate interactions between attributes. For example, the hypothesis of statistical independence between attributes is represented by a model in which all interactions are set to zero. The comparison between the fit of G-DINA model with and without interactions constitutes a statistical test of the hypothesis of independent attributes.

Empirical study

This section describes three empirical analyses aimed at developing and validating the ENCUIST for its use in personality profiling. The first empirical analysis consists of investigating the factor structure of the ENCUIST. In the second analysis, we used the CDM technique to classify the subjects in personality profiles. We also investigated the predictive validity of the dichotomous attributes in relation to two exogenous variables that are relevant in the forensic context. The third analysis consists of fitting a structural equation model (SEM) to ENCUIST to investigate the similarities and differences between CDM and SEM.

General Procedure

The three data analyses involve the use of the same empirical sample and instruments. These are described below.

Participants. The sample consisted of 516 university students, of whom 37.2% were men and 65.3% had completed postgraduate studies. The mean age was 32.8 years (*SD* = 11.5). Some characteristics of the sample that might be related to felonious behavior are as follows, 29.7% of the sample reported sometimes having been involved in fights, 3.1% had been incarcerated at some time, and 4.1% had a close relative who had been incarcerated. With regard to their relationships and capacity to establish links, 71.9% of the sample reported that they usually had stable couple relations (more than 2 years), and 12.2% reported that they changed friends easily. Lastly, we explored a victimological variable and found that 28.9% stated that they had suffered some kind of physical, sexual, or psychological abuse.

Instruments. As mentioned in the introduction, the ENCUIST is intended to measure the traits: E, N, CU, and EA. Additionally, we have investigated the criterion-related validity of ENCUIS scores in relation to two variables, BIS and BAS.

To assess the E and N traits, we used the 12-item E and N subscales of the five-factor Revised Neo Personality Inventory (NEO-FFI; Costa & McCrae, 2008). The response format ranges from 1 (*strongly disagree*) to 5 (*strongly agree*). With regard to the E scale, some examples of items are: "*I enjoy parties where there are many people*" and "*I like to be where the action is*". The E scale yields a Cronbach alpha of .83. Regarding the N scale, some examples of items are "*I often feel inferior to others*" and "*Sometimes it seems that I'm not worth a thing*" (Cronbach alpha of .85).

To assess the CU trait, we used the *Inventory of Callous-Unemotional Traits* (*ICU*, Roose, Bitjttebier, Decoene, Claes, & Frick, 2009), a 24-item self-report with 4 response options ranging from 1 (*not true*) to 4 (*definitely true*). The test is composed of three subscales: *Uncaring*, *Callousness*, and *Unemotional* (Cronbach alpha of .59, .67, and .86, respectively). Some examples of items are "*I don't care who I hurt as long as I get what I want*" and "*I don't care if I get into trouble or problems*".

To assess the EA trait, we used anger expression and anger control of the State-Trait Anger Expression Inventory–2 (STAXI–2; Spielberger, 1988). Specifically, the subscale of *External expression of anger* evaluates how the subject reacts when feeling furious. It is made up of 6 items with 4 response options (A: *almost never*, B: *sometimes*, C: *often*, and D: *almost always*), and it has a Cronbach alpha of .73. Some examples of items are: "*I express my anger*" and "*I make ironic comments about others*".

The BIS/BAS test was used to study the criterion-related validity of the ENCUIST scores (Carver & White, 1994). The BIS/BAS is made up of 20 items with 4 response options ranging from 1 (*strongly disagree*) to 4 (*strongly agree*). It contains two subscales: the BIS scale with 6 items and the BAS scale with 13 items (α = .74 and .77, respectively). Some examples of items are: "*When I think that something unpleasant will happen, I usually feel very upset and restless*" and "*In comparison with my friends, I have very few fears*".

Procedure. We drew up an instrument consisting of all the items described above, 74 in total, and this was administered online by means of the *Google Docs* application. This application generates a link that is

https://doi.org/10.1017/sjp.2018.62 Published online by Cambridge University Press

distributed to all study participants so that they can complete the questionnaire from their own computers. We informed the participants that their data were anonymous and confidential. All statistical analyses were run using the R language, version 3.4.1 (R Core Team, 2013).

Analysis 1: Factor structure of the ENCUIST

The computer program SPSS–22 was used to run factor analysis. First, we selected the items of the ENCUIST test with the best psychometric properties according to item-test correlation and reliability when the item was deleted. We selected the four best items for each one of the traits. The selected items for the E, N, CU and EA traits had Cronbach's alpha coefficients of .70, .78, .87, and .71, respectively. These 16 items were then used in subsequent analyses. The BIS and the BAS scales had Cronbach's alpha coefficients of .71 and .77, respectively.

Afterwards, we studied the factorial structure of the selected items by means of exploratory factor analysis. Given that our data are ordinal, a weighted least-squares estimation algorithm was applied following the recommendation of Flora and Curran (2004). The criteria suggested by Fabrigar and Wegener (2011) were used to determine the optimal number of factors; specifically, parallel analysis and goodness-of-fit statistics.

A parallel analysis (Horn, 1965) was performed to determine the number of dimensions measured by the 16 selected items. Parallel analysis compares the eigenvalues of the observed data matrix with the eigenvalues of K simulated data matrices of the same size as the observed data. The mean and the 95th percentile of simulated eigenvalues were compared to the observed eigenvalue. As is well known, the number of retained factors is the number of observed eigenvalues that are higher than the simulated eigenvalues. The results are shown in Table 1 and indicate that four factors suffice to explain the data.

The second method we used for selecting the number of factors is the chi-square for the hypothesis of exact fit. The chi-square statistic was computed to compare fit against a saturated model (χ^2) and to compare the goodness of fit of models with an increasing number of factors ($\Delta\chi^2$). The incremental chi-square, $\Delta\chi^2$, is computed between the model with one factor and the model with two factors, then between models with two and three factors, etc. This process is repeated until a non-significant chi-square is found. Finally, we computed the close fit statistics RMSEA, CFI and SRMR. The results appear in Table 2. The chi-square indicates that at least five factors are needed to explain the data (models with more than five factors were not estimated). All in all we concluded that a model with four factors is a sufficient, although not exact, representation of the data for two reasons, the result of the parallel analysis and the measures of close fit (CFI > .95 and RMSEA < .08). Moreover, from a theoretical viewpoint the four-factor model is better grounded in the personality literature, and theoretical justification can also be considered as another indicator of the model's goodness of fit (Reise, Widaman, & Pugh, 1993).

Factor loadings for the four-factor model were rotated to assist with their interpretation. Given that the factors could be related to each other, we carried out oblique

Table 1. *Results of Parallel Analysis for the ENCUIST*

Factor	Actual eigenvalue	Mean eigenvalue	95th Percentile eigenvalue
1	4.0190	1.3125	1.3797
2	2.7996	1.2473	1.2929
3	1.6750	1.1962	1.2357
4	1.2853	1.1528	1.1857
5	0.8141	1.1130	1.1446
6	0.7187	1.0759	1.1037
7	0.6945	1.0414	1.0685
8	0.6571	1.0081	1.0349
9	0.5669	0.9755	1.0021
10	0.5116	0.9102	0.9369
11	0.5116	0.9102	0.9369
12	0.4415	0.8776	0.9057
13	0.4300	0.8438	0.8733
14	0.3236	0.8082	0.8388
15	0.3040	0.7704	0.8031
16	0.2291	0.7238	0.7646

Table 2. *Goodness of Fit for Factorial Models between 1 and 5 Factors for the ENCUIST*

Factors	χ^2	*df*	(p-χ^2)	χ^2/df	RMSEA	CFI	SRMR	$\Delta\chi^2$	Δdf	p-$\Delta\chi^2$
1	3,849.03	104	< .001	37.01	.264	.683	.181			
2	1,461.35	89	< .001	16.42	.173	.884	.096	2387.68	15	< .001
3	685.62	75	< .001	9.14	.125	.948	.059	775.73	14	< .001
4	217.63	62	< .001	3.51	.069	.987	.028	467.99	13	< .001
5	112.92	50	< .001	2.26	.049	.995	.021	104.71	12	< .001

https://doi.org/10.1017/sjp.2018.62 Published online by Cambridge University Press

Table 3. *Oblimin and Procrustes Rotation for the Five-Factor Model*

	Oblimin pattern matrix				Procrustes rotation results				
Items	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2	Factor 3	Factor 4	Item congruence
E_1	**.66**	.08	.14	–.04	**.66**	.04	.15	–.05	.97
E_2	**.69**	–.09	–.03	.05	**.68**	–.12	–.02	.04	.98
E_3	**.65**	–.24	–.01	.01	**.63**	–.28	–.01	.02	.91
E_4	**.55**	.06	.002	.04	**.55**	.03	.01	.03	1.00
N_1	–.06	**.69**	–.05	.04	–.02	**.69**	–.04	.00	1.00
N_2	.02	**.92**	–.05	–.01	.08	**.92**	–.03	–.07	.99
N_3	–.31	**.58**	–.04	.08	–.27	**.60**	–.03	.05	.91
N_4	–.15	**.46**	.10	.19	–.12	**.47**	.10	.17	.90
ICU_1	.03	.06	**.85**	–.01	.02	.04	**.85**	.00	1.00
ICU_2	–.05	–.10	**.87**	.03	–.07	–.12	**.87**	.04	.99
ICU_3	.24	.25	**.71**	.02	.25	.22	**.71**	.01	.91
ICU_4	–.05	–.13	**.91**	–.02	–.07	–.15	**.90**	.00	.98
STAXI_1	–.01	–.03	.10	**.59**	.00	.00	.10	**.60**	.99
STAXI_2	.03	–.02	–.10	**.59**	.04	.02	–.11	**.59**	.98
STAXI_3	.05	.03	–.04	**.73**	.06	.07	–.04	**.73**	.99
STAXI_4	–.06	.02	.11	**.67**	–.05	.06	.10	**.67**	.98
Factor congruence					.95	.93	.98	.98	.96
Alpha for the factor					.695	.778	.871	.708	

Note: E refers to the items of Extraversion; N, to Neuroticism; ICU, to Callous Unemotional; and STAXI, to Expression of Anger. The boldfaced values are the ones with target 1 in the Procrustes rotation.

rotation (*oblimin* direct). As seen in the left section of Table 3, the E items are grouped in the first factor, the N items in the second factor, the CU items in the third, and the EA items in the fourth factor. The correlations between the different factor pairs show the greatest relation between N and EA (r = .33), E and CU (r = .37), and E and N (r = –.30). In contrast, lower relations were found between EA and E (r = –.02), and between EA and CU (r = .07). In addition to the oblimin rotation, we carried out a Procrustes rotation, which allows modifying the factor solution so that it conforms as much as possible to a preset theoretical pattern based on the assumption that the items are one-dimensional and measure only the attribute for which they were designed. Adaptation to the theoretical pattern was measured by means of similarity coefficients. The results appear on the right side of Table 3. As shown in the column of item congruence, all the values of the congruence coefficients are above .90; indicating a good congruence between the theoretical matrix and the rotated matrix, which supports interpreting the four-factor as E, N, ICU and STAXI.

Table 4. *Q-matrix used for the G-DINA Model*

	E	N	CU	EA
E_1	1	0	0	0
E_2	1	0	0	0
E_3	1	0	0	0
E_4	1	0	**1**	**1**
N_1	0	1	0	0
N_2	0	1	0	0
N_3	**1**	1	0	0
N_4	0	1	0	**1**
ICU_1	0	0	1	0
ICU_2	0	0	1	0
ICU_3	0	0	1	0
ICU_4	0	0	1	0
STAXI_1	0	0	0	1
STAXI_2	0	0	**1**	1
STAXI_3	0	0	0	1
STAXI_4	0	0	0	1

Note: Boldface means that the value is zero in the theoretical Q-matrix and it is changed to 1 by the Q-validation procedure.

Analysis 2: Application of Cognitive Diagnostic Models

Once the items for the ENCUIST were selected, we conducted a cognitive diagnostic analysis. The four factors of the model fitted in Analysis 1 are represented by four binary attributes in the G-DINA model. In these analyses, we investigated the relations between items and attributes, the associations among the four attributes and, finally, the linear regressions for BIS and BAS on the attributes to obtain evidence about criterion-related validity of the ENCUIST scores. To apply G-DINA the responses to the ENCUIST were dichotomized, response categories indicating any level of disagreement

https://doi.org/10.1017/sjp.2018.62 Published online by Cambridge University Press

Table 5. *Goodness of Fit for the Theoretical and the Empirical G-DINA Models*

Model	AIC	BIC	log *Lik.*	#Params	χ^2	*df*	*p*
Provisional	8,314.34	8,496.92	–4,114.17	43			
Empirical	8,303.46	8,537.00	–4,096.73	55	34.88	12	< .001

Note: The Q-matrix for the theoretical model represents the psychological attributed supposedly measured by the items. The empirical model uses the Q-matrix suggested by the modification index.

Table 6. *Parameters of the GDINA Model*

Item	Parameter	*Estimate*	*SE*	*z*	*p*	RMSEA
E_1	Baseline	–0.735	0.160	–4.580	< .001	.082
	E	2.684	0.229	11.683	< .001	
E_2	Baseline	–1.281	0.182	–7.035	< .001	.052
	E	3.464	0.256	13.528	< .001	
E_3	Baseline	–1.816	0.216	–8.387	< .001	.079
	E	3.081	0.253	12.172	< .001	
E_4	Baseline	0.009	0.203	0.045	.964	.040
	E	1.109	0.340	3.258	.001	
	CU	–0.708	0.431	–1.641	.101	
	EA	–0.461	0.449	–1.027	.304	
	E–CU	0.961	0.545	1.762	.078	
	E-EA	1.783	0.836	2.132	.033	
	CU-EA	2.195	0.752	2.919	.004	
	E-CU-EA	–3.117	1.105	–2.821	.005	
N_1	Baseline	–2.501	0.181	–13.804	< .001	.038
	N	3.833	0.328	11.686	< .001	
N_2	Baseline	–3.133	0.239	–13.074	< .001	.045
	N	4.241	0.351	12.060	< .001	
N_3	Baseline	–1.966	0.279	–7.030	< .001	.068
	E	–1.574	0.440	–3.578	< .001	
	N	2.798	0.398	7.018	< .001	
	E-N	–0.331	0.714	–0.464	.642	
N_4	Baseline	–2.060	0.174	–11.836	< .001	.076
	N	1.980	0.366	5.406	< .001	
	EA	0.522	0.308	1.695	.090	
	N-EA	0.758	0.576	1.317	.188	
ICU_1	Baseline	–1.644	0.180	–9.123	< .001	.048
	ICU	3.845	0.266	14.429	< .001	
ICU_2	Baseline	–1.563	0.175	–8.912	< .001	.057
	ICU	3.743	0.262	14.282	< .001	
ICU_3	Baseline	–0.709	0.141	–5.024	< .001	.080
	ICU	3.163	0.259	12.186	< .001	
ICU_4	Baseline	–1.457	0.169	–8.593	< .001	.072
	ICU	4.061	0.287	14.121	< .001	
STAXI_1	Baseline	–1.321	0.127	–10.335	< .001	.063
	EA	2.584	0.235	10.961	< .001	
STAXI_2	Baseline	–2.866	0.341	–8.403	< .001	.038
	CU	0.072	0.457	0.159	.874	
	EA	3.546	0.439	8.067	.000	
	CU-EA	–1.077	0.575	–1.872	.061	
STAXI_3	Baseline	–2.823	0.226	–12.453	< .001	.055
	EA	3.211	0.281	11.401	< .001	
STAXI_4	Baseline	–1.772	0.148	–11.976	< .001	.077
	EA	2.813	0.238	11.796	< .001	

https://doi.org/10.1017/sjp.2018.62 Published online by Cambridge University Press

were assigned the code 0, and the responses indicating agreement received the code 1. The neutral category received the code 0 to equilibrate response proportions in the dichotomized data, because in our sample categories four and five are more frequent than categories one and two.

Empirical validation of the Q-matrix. The construction of the Q-matrix is typically performed by domain experts and it is to a large extent a subjective process. Recently, de la Torre and Chiu (2016) proposed a general method for empirical validation of the Q-matrix. This method requires a provisional Q-matrix based on expert judgements, and consists of computing a discrimination index that can be used to identify and correct misspecified entries of the Q-matrix. The Q-matrix empirical validation for the present study was conducted using the Qval function of the GDINA R package (Ma & de la Torre, 2018).

The provisional Q-matrix for the present analysis contains ones in the positions that correspond to the attribute that theoretically is measured by the item and zeroes in the other entries, in concordance to a simple structure.

Table 4 contains the Q-matrix suggested by the empirical validation procedure. Only five elements of the provisional Q-matrix were changed by the empirical validation procedure. The Q-matrix in Table 4 was used in all subsequent G-DINA analyses.

The goodness of fit for G-DINA models based on the provisional (theoretical) and empirically validated Q-matrix appear in Table 5, including the likelihood-ratio chi-square and distance measures. The result is that the difference between models is significant, and thus the empirically validated Q-matrix is used in the subsequent analysis.

Results for the G-DINA model

The G-DINA model based on the empirically validated Q-matrix was estimated using the *CDM* package (George, Robitzsch, Kiefer, Groß, & Ünlü, 2016; Ravand & Robizsch, 2015; Robitzsch, Kiefer, George, & Uenlue, 2017). Table 6 presents the information about the estimates of Equation (1), including the values of the estimators, their standard errors, and the *z* statistic, which contrasts the hypothesis that the true value of the parameter is zero. As an example of the interpretation of these results, we can use item N_3. This item measures the attributes E and N, so the parameters shown in Table 6 are the baseline (δ_{j0} in Equation 1), the weight of each attribute on the item (δ_{j1} and δ_{j2}), and the interaction of the two attributes within the item (δ_{j12}). Item N_3 has a more extreme *z* in attribute N, indicating that neuroticism is the dominant attribute for this item. In contrast, the *z*-value for the E-N interaction is not significant. In addition, in Table 6, we can observe that none of the interaction parameters between pairs of attributes are significantly different from zero.

Table 7 shows the probabilities of the 16 personality profiles, that is, the possible classifications over the four dichotomous attributes. These probabilities are the parameters of Equation (2). As can be observed, the most likely profile (1010) contains subjects who claim that they are extroverts, emotionally stable, unemotional, and who do not express anger (30.1% of the sample).

Associations between attributes. We used tetrachoric correlations to analyze the relations between the binary values representing attributes. The tetrachoric correlations showed that scores in *E* are associated with N (–.39) and with CU (.51); and have hardly any relation with EA (.02). Regarding N, there is a negative relation between N and CU (–.33), and a positive relation between N and EA (.39). Finally, there is no significant relation between CU and EA (.003).

We also applied log-linear models using the *glm* function of R to investigate patterns of associations among the binary attributes. The chi-square is 107.91 (df = 11) for the model with no interactions and $\chi^2 = 0.0002$ ($df = 5$) for the model with pairwise interactions. Thus the results indicate that pairwise interactions suffice to explain the distribution of attributes. Table 8 shows the estimated log-linear parameters, ω, of Equation 3 for the model with pairwise interactions.

Criterion-related validity of the ENCUIST. Linear regressions were computed to provide evidence about the predictive validity of the ENCUIST taxonomies in relation to BIS and BAS. The independent variables

Table 7. *Estimated Probabilities for the 16 Personality Profiles*

Profile	Probability
0000	.129
0001	.146
0010	.044
0011	.007
0100	.059
0101	.300
0110	.014
0111	.010
1000	.025
1001	.041
1010	.391
1011	.092
1100	.016
1101	.122
1110	.018
1111	.019

Note: The profile is the binary classification in the E, N, ICU, and STAXI attributes.

https://doi.org/10.1017/sjp.2018.62 Published online by Cambridge University Press

Table 8. *Results of the Log-Linear Model to the Cross-Classification of Attributes*

Parameter	*Estimate*	*SE*	*z*	*p*
Intercept	4.203	.116	36.256	< .001
E	0.116	.153	0.756	.450
N	–1.102	.210	–5.257	< .001
ICU	–0.801	.190	–4.219	< .001
STAXI	–1.656	.234	–7.081	< .001
E–N	–1.902	.300	–6.339	< .001
E-ICU	1.524	.212	7.182	< .001
E-STAXI	0.389	.252	1.544	.123
N-ICU	–0.381	.289	–1.320	.187
N-STAXI	1.555	.279	5.576	< .001
ICU-STAXI	0.364	.222	1.639	.101

for the linear regressions are the binary scores representing the attributes, including the interactions of any order among attributes. According to the hypothesis about the BAS, we expected to find a positive relationship between this variable and trait E, and, regarding BIS, we expected that high scores on BIS would be related to high scores on EA, whereas hypo-activity on BIS would be related to high scores on CU.

The results are significant, but there is still room for improvement. Regarding the BIS, the adjusted R^2 is .070, $F_{15,\,500} = 3.59$, $p < .001$. The results for the BAS are adjusted $R^2 = .106$, $F_{15,\,500} = 5.07$, $p < .001$. Table 9 shows the estimated parameters. Only some main effects of the attributes reached statistical significance; no evidence was found about the relevance of interactions between attributes in explaining BIS and BAS. The coefficients of N and STAXI on the BIS equation are significant. The significant predictors for BAS are E and ICU. These results conform to the expected relations among variables.

Analysis 3: Fitting a SEM model

The SEM model was fitted to the ENCUIST data to illustrate the differences and similarities between SEM and CDM. The SEM model was estimated using the R package *lavaan* (Rosseel, 2012). The four attributes, E, N, ICU, STAXI, are the factors of the SEM model; each factor is measured by the same observable variables as in the CDM. The linear regression between the criterion variables (BIS and BAS) and the factors were included in the SEM model. The SEM was estimated from the original, non-dichotomized, responses to the ENCUIST.

Table 10 contains the estimated item parameters for the SEM. Table 10 is the SEM analog of Table 6. The main difference between SEM and CDM is that the later model provides an estimate of the interactions between factors (attributes in the CDM terminology) within the items, see for example item E_4, whereas these interactions are not included in SEM.

The estimated factor correlations and their standard errors appear in Table 11. This table is the SEM analog of Table 8, which contains the associations of attributes for the CDM model. The SEM approach assumes multivariate normality for the factors and thus, the only type of association between factors is linear correlation. In contrast, the CDM approach allows estimation of interactions of any order among attributes by fitting a log-linear model to the binary variables representing the scores in the attributes. In the present investigation we only

Table 9. *Linear regression of the BIS and the BAS over the attributes of the CDM and their interactions*

	Dependent variable: BIS				Dependent variable: BAS			
Effect	*Estimate*	*SE*	*z*	*p*	*Estimate*	*SE*	*z*	*p*
Intercept	18.676	0.317	50.003	< .001	39.191	0.545	71.866	< .001
E	–0.207	0.451	–0.458	.647	2.611	0.777	3.361	.000
N	2.074	0.664	3.123	.002	–1.541	1.143	–1.347	.178
ICU	–0.110	0.533	–0.204	.838	2.457	0.918	2.675	.007
STAXI	1.390	0.745	1.867	.062	1.275	1.282	0.994	.320
E-N	–0.981	0.984	–0.996	.320	0.050	1.696	0.030	.976
E-ICU	0.329	0.670	0.492	.623	–1.268	1.149	–1.104	.269
N-ICU	–0.356	1.264	–0.281	.779	–0.250	2.178	–0.115	.908
E-STAXI	–0.260	0.940	–0.276	.782	1.221	1.620	0.754	.451
N-STAXI	–1.724	1.209	–1.425	.155	1.324	2.083	0.636	.525
ICU-STAXI	–0.672	1.107	–0.607	.544	–1.138	1.906	–0.597	.550
E-N-ICU	0.032	1.571	0.020	.983	0.832	2.706	0.308	.758
E-N-STAXI	2.656	1.752	1.516	.130	1.616	3.018	0.535	.592
E-ICU-STAXI	–0.613	1.315	–0.466	.641	–1.694	2.266	–0.748	.454
N-ICU-STAXI	2.250	1.873	1.201	.230	0.093	3.226	0.029	.977
E-N-ICU-STAXI	–1.886	2.634	–0.716	.475	1.019	4.537	0.225	.822

https://doi.org/10.1017/sjp.2018.62 Published online by Cambridge University Press

Table 10. *Parameters of the Measurement Sub-Model for the SEM Model*

Item	Parameter	*Estimate*	*SE*	*z*	*p*
E_1	E	.541	.040	13.653	< .001
	Error variance	.468	.036	13.007	< .001
E_2	E	.710	.051	13.994	< .001
	Error variance	.752	.059	12.782	< .001
E_3	E	.855	.055	15.616	< .001
	Error variance	.779	.068	11.445	< .001
E_4	E	.488	.053	9.141	< .001
	CU	–.066	.050	–1.316	.188
	EA	.115	.047	2.476	.013
	Error variance	.683	.048	13.019	< .001
N_1	N	.847	.049	17.153	< .001
	Error variance	.656	.055	11.939	< .001
N_2	N	.981	.049	20.180	< .001
	Error variance	.449	.054	8.266	< .001
N_3	E	–.306	.053	–5.833	< .001
	N	.640	.053	12.139	< .001
	Error variance	.626	.048	13.019	< .001
N_4	N	.524	.054	9.634	< .001
	EA	.194	.056	3.439	.001
	Error variance	.925	.063	14.772	< .001
ICU_1	ICU	.802	.039	20.611	< .001
	Error variance	.387	.031	12.593	< .001
ICU_2	ICU	.860	.039	22.320	< .001
	Error variance	.321	.029	11.135	< .001
ICU_3	ICU	.635	.039	16.379	< .001
	Error variance	.501	.035	14.434	< .001
ICU_4	ICU	.880	.037	23.811	< .001
	Error variance	.243	.026	9.379	< .001
STAXI_1	EA	.428	.038	11.122	< .001
	Error variance	.443	.037	12.065	< .001
STAXI_2	CU	–.094	.037	–2.577	.010
	EA	.542	.041	13.205	< .001
	Error variance	.458	.033	13.691	< .001
STAXI_3	EA	.591	.041	14.257	< .001
	Error variance	.430	.039	11.169	< .001
STAXI_4	EA	.542	.041	13.377	< .001
	Error variance	.443	.037	12.076	< .001

found pairwise association between attributes. Thus, correlation suffices to explain the association pattern.

Finally, Table 12 contains the regression sub-model of SEM, which consists of the linear regression between BIS and BAS over the factors measured by ENCUIST. This is the SEM analog of Table 9. The pattern of results for CDM and SEM concur with some exceptions. Similar to CDM, the BIS variable is predicted by the *N* factor and the BAS variable is predicted by E under SEM. However, a significant relation between ICU and BAS is found for CDM and is missing for SEM. The contrary occurs for the relation between STAXI and BAS.

Table 11. *Factor Variances and Covariances for the SEM Model*

	E	N	ICU	STAXI
E	1	.051	.048	.062
N	–.434	1	.050	.054
ICU	.439	–.240	1	.057
STAXI	–.026	.324	.121	1

Note: Factor correlations appear below the diagonal. Standard errors appear above the diagonal.

Discussion

Our aim was to develop and validate the ENCUIST questionnaire for personality profiling in a forensic context. The ENCUIST tests measures the traits of Extroversion (E), Neuroticism (N), Callous-Unemotional (CU) and

https://doi.org/10.1017/sjp.2018.62 Published online by Cambridge University Press

Table 12. *Parameters of the Regression Sub-Model for the SEM Model*

Dependent variable	Factor	*Estimate*	*SE*	*z*	*p*
BIS					
	E	.001	0.173	0.004	.997
	N	.947	0.166	5.689	< .001
	ICU	.016	0.141	0.111	.912
	STAXI	.235	0.150	1.572	.116
	Error variance	6.216	.402	15.475	< .001
BAS					
	E	2.062	0.305	6.769	< .001
	N	.274	0.290	0.945	.345
	ICU	.028	0.251	0.112	.911
	STAXI	.806	0.263	3.062	.002
	Error variance	17.968	1.234	14.564	< .001

Expression of Anger (EA) using a limited number of items. In the forensic context, the questionnaires have to be short, because a rapid initial evaluation of the subject is needed prior to more in-depth interviews. CDMs provide a means to classify individuals as high or low in relevant psychological attributes. Exploratory factor analysis was used to select the four items that best measure each of the traits, which resulted in a final questionnaire of 16 items.

The methodological tool used to elaborate personality profiles from responses to the ENCUIST is the G-DINA model, taken from the field of CDMs. The G-DINA model provides a unique profile for each subject consisting of a classification on binary personality attributes.

Linear regressions were estimated to evaluate the predictive validity of the personality profiles with respect to the BIS and BAS variables. The results support the proposed hypothesis about the relation between these variables and the attributes measured by the ENCUIST, although the effect sizes were low. The variable that best predicts BAS scores is E. The BAS underlies the dimension of impulsivity and is activated by signals of reward and is thereby related to extroverted subjects (Newman et al., 2005). As expected, the variable N is the one that best predicts the BIS score. The BIS is a system that has been linked in diverse studies of anxiety and negative moods (Carver & White, 1994). Given that the BIS variable has been shown to be related to emotional instability, it is consistent that it is also related to the variable EA (Lievaart et al., 2016). In contrast, we did not find any relation between low scores on the BIS and high scores on CU in the regression analysis results, but we did find a relation when analyzing the tetrachoric correlations between the different attributes. In these analyses, we obtained a negative correlation between the variables N and CU. These results indicate that emotionally stable people with low anxiety scores typically obtain higher scores on CU.

The results of the CDM were compared to a SEM model to illustrate the differences between the two approaches. In the present investigation the CDM makes more sense because the psychological attributes are to be treated as binary variables in the forensic context. The CDM gives dichotomized attribute scores based on optimizing the statistical goodness of fit to the current data, thus the binary classification of CDM is optimal in a statistical sense. On the other hand, latent trait models (i.e., factor analysis and SEM) provide factor scores along a continuous scale that would have to be dichotomized a-posteriori using arbitrary cut-points to classify individuals. The cut-point is based on the scale of the latent factor or its distribution in the population, but it is not intended to optimize a statistical fit measure.

The present paper illustrates that the CDM approach is more flexible than the SEM approach in investigating interactions of attributes within the items and the interactions between the attributes. SEM models are based on linear relations between observed variables and factors, and between factors. In consequence, associations between variables reduce to correlations. In contrast, CDM is based on binary variables and thus, log-linear models (with interactions of any order) can be used to investigate the associations between observed variables and attributes, and among the attributes themselves. In this investigation the conclusions from CDM and SEM showed similar general patterns because the triple or higher order interactions happened to be non-significant. The advantages of CDMs over SEM models should be more prominent in those applications in which there are interactions of attributes within items, in the distribution of attributes in the population of individuals, or in the regressions of criterion variables over attributes.

The present investigation is only a first step in the development of the ENCUIST that will be continued by gathering data from forensic populations. The main

https://doi.org/10.1017/sjp.2018.62 Published online by Cambridge University Press

limitation of the present paper is the use of a convenience sample to calibrate ENCUIST. We did so because the large samples needed to apply the test-construction and statistical procedures are difficult to obtain in a forensic context. This sample is still informative because the traits measured by ENCUIST are general personality traits that are not specific for the forensic population. Notwithstanding, ENCUIST still has to be validated using data from people involved in judicial processes; including witnesses, defendants, or those prison sentences. Particularly important are data from persons who have engaged in actual criminal behavior. For example, one identifying characteristic for inclusion in such a sample would be whether an individual has ever been arrested and charged with a serious crime, and then whether or not a finding of guilt has been determined. Future research should focus on validating ENCUIT using data from forensic populations.

References

Barbour K. A., Eckhardt C. I., Davison G. C., & Kassinove H. (1998). The experience and expression of anger in maritally violent and maritally discordant-nonviolent men. *Behavior Therapy*, *29*, 173–191. https://doi.org/10.1016/S0005-7894(98)80001-4

Baron K. G., Smith T. W., Butner J., Nealey-Moore J., Hawkins M. W., & Uchino B. N. (2007). Hostility, anger, and marital adjustment: Concurrent and prospective associations with psychosocial vulnerability. *Journal of Behavioral Medicine*, *30*, 1–10. https://doi.org/10.1007/s10865-006-9086-z

Blair R. J. R., Colledge E., Murray L., & Mitchell D. G. (2001). A selective impairment in the processing of sad and fearful expressions in children with psychopathic tendencies. *Journal of Abnormal Child Psychology*, *29*, 491–498. https://doi.org/10.1023/A:1012225108281

Blair R. J. R., Mitchell D. G. V., Peschardt K. S., Colledge E., Leonard R. A., Shine J. H., … Perret D. I. (2004). Reduced sensitivity to other's fearful expressions in psychopathic individuals. *Personality and Individual Differences*, *37*, 1111–1122. https://doi.org/10.1016/j.paid.2003.10.008

Carver C. S., & White T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, *67*, 319–333. https://doi.org/10.1037/0022-3514.67.2.319

Chen J., & de la Torre J. (2013). A general cognitive diagnostic model for expert-defined polytomous attributes. *Applied Psychological Measurement*, *37*, 419–437. https://doi.org/10.1177/0146621613479818

Corr P. J. (2010). The psychoticism–psychopathy continuum: A neuropsychological model of core deficits. *Personality and Individual Differences*, *48*, 695–703. https://doi.org/10.1016/j.paid.2009.12.023

Costa P., & McCrae R. (2008). Revised Neo Personality Inventory. In G. J. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *The SAGE handbook of personality theory and assessment* (Vol. 2). Thousand Oaks, CA: SAGE.

de la Torre J. (2009). DINA model and parameter estimation: A didactic. *Journal of Educational and Behavioral Statistics*, *34*, 115–130. https://doi.org/10.3102/1076998607309474

de la Torre J. (2011). The generalized DINA model framework. *Psychometrika*, *76*, 179–199. https://doi.org/10.1007/s11336-011-9207-7

de la Torre J., & Chiu C.-Y. (2016). A general method of empirical Q-matrix validation. *Psychometrika*, *81*(2), 253–272. https://doi.org/10.1007/s11336-015-9467-8

de la Torre J., van der Ark L. A., & Rossi G. (2015). Analysis of clinical data from cognitive diagnostic modeling framework. *Measurement and Evaluation in Counseling and Development*, *48*, 1–16. https://doi.org/10.1177/0748175615569110

Embretson S. (1984). A general latent trait model for response processes. *Psychometrika*, *49*, 175–186. https://doi.org/10.1007/BF02294171

Eysenck H. (1952). *The scientific study of personality*. Oxford, UK: Macmillan.

Fabrigar L. R., & Wegener D. T. (2011). *Exploratory factor analysis*. New York, NY: Oxford University Press.

Flora D. B., & Curran P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, *9*, 466–491. https://doi.org/10.1037/1082-989X.9.4.466

Frick P. J., & White S. F. (2008). Research review: The importance of callous-unemotional traits for developmental models of aggressive and antisocial behavior. *Journal of Child Psychology and Psychiatry*, *49*, 359–375. https://doi.org/10.1111/j.1469-7610.2007.01862.x

García P. E., Olea J., & de la Torre J. (2014). Application of cognitive diagnostic models to competency-based situational judgment tests. *Psicothema*, *26*, 372–377. https://doi.org/10.7334/psicothema2013.322

George A. C., Robitzsch A., Kiefer T., Groß J., & Ünlü A. (2016). The R Package CDM for cognitive diagnostic models. *Journal of Statistical Software*, *74*, 1–24. https://doi.org/10.18637/jss.v074.i02

Goldberg L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, *48*, 26–34. https://doi.org/10.1037/0003-066x.48.1.26

Gray J. A. (1970). The psychophysiological basis of introversion-extraversion. *Behavior Research & Therapy*, *8*, 249–266. https://doi.org/10.1016/0005-7967(70)90069-0

Haslam N., Holland E., & Kuppens P. (2012). Categories versus dimensions in personality and psychopathology: A quantitative review of taxometric research. *Psychological Medicine*, *42*, 903–920. https://doi.org/10.1017/S0033291711001966

Heym N., & Lawrence C. (2010). The role of Gray's revised RST in the P–psychopathy continuum: The relationships of psychoticism with a lack of fear and anxiety, and increased impulsivity. *Personality and Individual Differences*, *49*, 874–879. https://doi.org/10.1016/j.paid.2010.07.021

Horn J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*, 179–185. https://doi.org/10.1007/BF02289447

Jasisnki M. J., Lumley M. A., Latsch D., Schuster E., Kinner E., & Burns J. W. (2016). Assessing anger expression: Construct validity of three emotion

https://doi.org/10.1017/sjp.2018.62 Published online by Cambridge University Press

expression-related measures. *Journal of Personality Assessment*, *98*, 640–648. https://doi.org/10.1080/00223891.2016.1178650

Lievaart M., van der Veen F., Huijding J., Naeije L., Hovens J., & Franken I. (2016). Trait anger in relation to neural and behavioral correlates of response inhibition and error-processing. *International Journal of Psychophysiology*, *99*, 40–47. https://doi.org/10.1016/j.ijpsycho.2015.12.001

Ma W., & de la Torre J. (2018). *GDINA: The generalized DINA model framework, R package version 2.0*. [Computer software]. Retrieved from https://CRAN.R-project.org/package=GDINA

Maris E. (1999). Estimating multiple classification latent class models. *Psychometrika*, *64*, 187–212. https://doi.org/10.1007/BF02294535

Matthews G., & Gilliland K. (1999). The personality theories of H. J. Eysenck and J. A. Gray: A comparative review. *Personality and Individual Differences*, *26*, 583–626.

Meyer T. D., & Hofmann B. U. (2005). Assessing the dysregulation of the behavioral activation system: The hypomanic personality scale and the BIS-BAS scales. *Journal of Personality Assessment*, *85*, 318–324. https://doi.org/10.1207/s15327752jpa8503_08

Newman J. P., MacCoon D. G., Vaughn L. J., & Sadeh N. (2005). Validating a distinction between primary and secondary psychopathy with measures of Gray's BIS and BAS constructs. *Journal of Abnormal Psychology*, *114*, 319–323. https://doi.org/10.1037/0021-843X.114.2.319

R Core Team (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing [Computer software]. Vienna, Austria. Retrieved from http://www.R-project.org/

Ravand H., & Robitzsch A. (2015). Cognitive diagnostic modeling using R. *Practical Assessment, Research & Evaluation*, *20*, 1–12.

Reise S. P., Widaman K. F., & Pugh R. H. (1993). Confirmatory factor analysis and item response theory: two approaches for exploring measurement invariance. *Psychological Bulletin*, *114*, 552–566. https://doi.org/10.1037/0033-2909.114.3.552

Robitzsch A., Kiefer T., George A. C., & Uenlue A. (2017). *CDM: Cognitive diagnostic modeling. R package version 6.0–101*. [Computer software]. Retrieved from https://CRAN.R-project.org/package=CDM

Roose A., Bitjttebier P., Decoene S., Claes L., & Frick P. (2009). Assessing the affective features of psychopathy in adolescence: A further validation of the inventory of callous and unemotional traits. *Assessment*, *17*, 44–57. https://doi.org/10.1177/1073191109344153

Rosseel Y. (2012). Iavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48* (2), 1–36. https://doi.org/10.18637/jss.v048.i02

Smith T. W, Glazer K., Ruiz J. M., & Gallo L. C. (2004). Hostility, anger, aggressiveness, and coronary heart disease: An interpersonal perspective on personality, emotion, and health. *Journal of Personality*, *72*, 1217–1270. https://doi.org/10.1111/j.1467-6494.2004.00296.x

Sorrel M. A., Olea J., Abad F. J., de la Torre J., Aguado D., & Lievens F. (2016). Validity and reliability of situational judgement test scores: A new approach based on cognitive diagnostic models. *Organizational Research Methods*, *19*, 506–532. https://doi.org/10.1177/1094428116630065

Spielberger C. D. (1988). *Manual for the State-Trait Anger Expression Inventory*. Odessa, FL: Psychological Assessment Resources.

Tatsuoka K. (1985). A probabilistic model for diagnosing misconceptions in the pattern classification approach. *Journal of Educational Statistics*, *12*, 55–73. https://doi.org/10.2307/1164930

Webster G. D., DeWall C. N., Pond R. S., Jr., Deckman T., Jonason P. K., Le B. M., … Bator R. J. (2015). The brief aggression questionnaire: Structure, validity, reliability and generalizability. *Journal of Personality Assessment*, *97*, 638–649, https://doi.org/10.1080/00223891.2015.1044093

https://doi.org/10.1017/sjp.2018.62 Published online by Cambridge University Press