Assessing Empathy: Development of an Observational Instrument

Tom van Erven, Jan B. Dijkstra, Val Reed

It was not I that sinn'd the sin,
The ruthless body dragged me in;
Though long I strove courageously,
The body was too much for me.

(***Walt Whitman: The Singer in the Prison)*** [1]

In forensic psychiatry the phenomenon of empathy plays an important part. Upon its precise description and operationalisation depends judgement of the empathic capability of a patient as a basis for expert advice to the Ministry of Justice and to mental health tribunals. It also plays an important part in treatment. It may be unclear precisely what various experts understand by the term; and especially on which concrete behaviours or observations the estimate of empathic capability is based. During development of the Behavioural Status Index (**Van Erven, 1999**) (an observational instrument then consisting of five sub-scales - risk, insight, communication and social skills, work and recreation, self and family care - with a total of 120 items whose format is similar to that reproduced in **Table I**) the necessity for a more objective assessment of empathic capability of forensic patients became increasingly apparent. With help from Concept Mapping[2] a pilot observational instrument was developed to describe empathic capability as objectively as possible in terms of concrete, observable behaviour.

Empathy

The word itself is derived from Greek *em* and *patheia*, signifying 'with affect or feeling'. Though the term is regularly used, operationalising it is no easy matter. This is because it functions as a portmanteau-word, meaning different things to different people. This has led in turn to its becoming " … a term widely used and written about ... [whose] … meaning and application have become blurred." (**Wiseman, 1996**).

The *Groot Woordenboek* (**Van Dale, 1999)** defines empathy as : "the ability to realise the feelings of others". Other sources cite or use alternative definitions (for example, **Barrett-Lennard, 1962; Gould, 1990; Duan and Hill, 1996; Wiseman, 1996; Eslinger, 1998; Binder, 1999; Vanaerschot, 1999; Van Strien, 1999**). However, all such definitions describe empathy as having something to do with consciousness of - or the capacity to enter into, and possibly sympathise with - another person's feelings. These relatively abstract ideas need to be expressed more completely on an observable, behavioural level before it becomes possible to develop a useful, objective observational instrument for assessing empathy.

---

[1] **Holloway, E. ed (1938):** Walt Whitman: Complete Verse and Selected Prose (First Edition), pp 344-345. Written in 1869, this poignant poem occurs in the section of *Leaves of Grass* entitled *Autumn Rivulets.* London: The Nonesuch Press.

[2] Concept Mapping is a method by which complex and diffuse concepts can be clarified and made 'workable'. Concept Mapping has six phases: 1) Determining the ***focus and choice*** of the participants, 2) A ***"brainstorm"*** phase for generating items; 3) A ***structuring*** phase, during which all items are clustered and relative values assigned; 4) ***Statistical analysis***, aided by a specially developed computer program (**Talcott/Ncgv, 1995**); 5) ***Interpretation***; and 6) ***Implementation***.

Need for an observational instrument

Apart from the following general problems with self-assessment, it was initially obvious that the nature of the patient population in our forensic clinic - and in many other forensic institutions - makes self-assessment on a substantial scale impossible. This is true of the Eindhoven clinic, where some 60 percent of patients are from non-Dutch cultural and/or linguistic backgrounds; and at least 50 percent suffer from chronic psychoses. Experience also teaches that, even with patients who are able and prepared to complete self-ratings, a considerable number drop out at the repeated-measures stage. Therefore our chosen starting-point was the development of a research instrument by means of which individuals who know a patient well can develop a reliable profile of his/her empathic capability (**Spill, 1999**).

Limitations of unaccompanied self-assessment

There are numerous indications in the literature that, unless accompanied by cross-validational observations, self-report measures are subject to serious methodological limitations. In the late forties, **Cronbach (1968)** drew attention to the so-called "hello-goodbye" effect, in which patterns of responding displayed by subjects in clinical self-report instruments varied significantly according to whether the individual was concerned to "fake into" or "fake out of" therapeutic programmes. Similarly, where socially involved issues are at stake - as in sexual behaviour - commentaries would suggest that behavioural data obtained by self-report methods are inherently unreliable and invalid due to multiple sources of bias and to intentional misrepresentation (**Lewontin, 1995; Brody, 1995**).

More recently, studies of the impact of depressive mood on response patterns in self-report have examined the reliability of biographical and personality data obtained by such means. Strong effects of high neuroticism and low aim-relatedness (leading to reporting of more life events and greater stress levels) occurred in the reporting of personality data. In the same study workers noted that negative primary socialisation experiences were directly associated with increased reported neuroticism and stress levels (**Bühler, Haltenhof *et al,* 1999**). With no differentiation possible between "personality" and "depression" in such studies, there is a clear need for preventive psychotherapeutic intervention to avoid "vicious circles" between life-events and mood-dependent factors, monitored both subjectively (by self-report) and objectively (by clinical-behavioural observation techniques).

Bias in Self-Reports:

The influence of social desirability "sets" and predominant response "styles" (either negative or positive) in affecting self-report responses is extensively described. For example, in a study comparing self-report data on height and weight with directly measured data, it was found that subjects tend to over-report their height but under-report their weight (**Hart and Tomazic, 1999**). Error patterns of this type are attributable to individuals wanting to present a "better" physical image to others. An analogy with reporting of personal psychological states by forensic psychiatric patients anxious to be discharged may be readily drawn.

Memory Errors:

A further problem with self-assessment concerns the fallibility of human memory. **Schachter (1999)** draws attention to seven flaws which may influence such assessments:

- **Transience:** decreasing accessibility of information over time;

- **Attention deficit:** inattentive or shallow processing contributing to a weak memory for feelings or events;
- **Blocking:** the temporary inaccessibility of information stored in memory, whether or not such blocking is "motivated" by acceptable self-image or the desire to impress others;
- **Misattribution:** attributing a recollection or an idea to the wrong source;
- **Suggestibility:** "memories" implanted by others as a result of leading questions or perceived expectations (as in the therapeutic context);
- **Bias:** retrospective distortions and unconscious influences which are related to current knowledge and beliefs, however mistaken;
- **Persistence:** pathological remembrances, information or events that we cannot forget, even though we wish we could.

- any or all of which may play a role in reducing validity and reliability of self-assessment data.

Measurement of empathy
Instruments purporting to measure empathy can be classified into three groups (**Reynolds and Presley, 1988**). These are

- ***self-assessment scales***: for example, the La Monica Empathy Profile (**La Monica, 1980a**); and the Empathy Construct Rating Scale (ECRS) (**La Monica, 1981, 1994**);

- ***expert appraisal*** : for example, the Hogan Empathy Scale (**Hogan, 1969**); the Empathy Scale (therapist's version) (**Burns and Auerbach, 1998**); the Empathy Scoring List (**Van Strien, 1999**): and thirdly

- ***client-patient appraisal*** : for example, the Empathy Sub-scale or the Barratt-Lennard Relationship Inventory (BLRI) (**Barratt-Lennard, 1962**); and the Empathy Scale (patient's version) (**Burns and Auerbach, 1998**).

Within client-centered and psychoanalytic schools, empathy is considered as a necessary characteristic of the therapist. Consequently, empathy 'measurement' in such treatment contexts bears a particular relation to the measurement of therapist empathy (e.g., **Ickes, 1997**; **Vanaerschot, 1999**; **Van Strien, 1999**). In the current investigation no such specific therapeutic groups are included.

Empathy as a construct
According to **Duan and Hill (1996)** empathy can refer to three distinct, non-overlapping constructs. Empathy can be considered either as a personality trait; or as a general skill; or as a situation-specific, cognitive-affective 'condition'. The first two constructs imply that, either as a result of their nature or their developmental experience, some people are consistently more 'empathic' than others. The third construct implies that empathic conduct varies according to situation, independently of a person's supposed 'level' of empathy. This last model suggests a situation-specific, moment-to-moment experience of empathy, implying that empathy varies according to context. Examination of this last construct is currently restricted to descriptive investigations, since objective assessment of the construct appears problematic if not impossible.

Repeated observation

Important in a TBS treatment setting[3] is the question whether people with a restricted or absent empathic capability can *learn* to become empathic. This is clinically important because a defective or absent empathic capability is often closely related to an individual's pattern of offending. If empathy can only be acquired during early learning, then subsequent attempts to 'teach' it may result in feigned or non-genuine behaviour. The question is, whether it is practically possible to learn to play an empathic role with objectives other than simply to get one's own way. Again, though the capacity completely to 'live through' another person's experience may be partially or completely lacking, it may be possible to help improve the quality of an individual's interpersonal contacts. If so, this might be confirmed by his/her altered conduct until, as a result of appropriate intervention, the probability of achieving more skilled empathic responses is increased. An effective observational instrument would make it possible systematically to record whether there were any signs of change in an individual's empathic conduct during a series of repeated measures, and if so in which direction. This would in turn make it clear whether an offender can *become* genuinely empathic; and ultimately what influence this will have on recidivism rates for offenders.

Concept Mapping

Concept Mapping is a well-tried method of determining evaluative criteria (**Trochim, 1989; Derks and Mulder, 1993**); carrying out task analyses (**Department of Judicial Planning, 1998**); generating quality criteria (**Coumans et al, 1994; Boevink and Wolf, 1998**), developing quality assessment instruments (**Nijssen et al, 1999**) and dealing with other complex problems (**De Boer, 1997**). For these purposes the program makes use of a computer-based analytic procedure (**Talcott bv./Ncgv, 1995**).

To facilitate development of a pilot research version of an empathy observational instrument, the following concept mapping procedure was carried out:

- A group of twelve expert psychologists and psychiatrists used the brainstorming phase to generate a total of 55 responses in answer to the incomplete question: "For me, some concrete indications of a patient's empathic capability include ... ";

- These responses were prioritised and clustered by twelve psychologists and ten psychiatrists (six women and sixteen men).

- With the help of the Concept Mapping program 'Ariadne' (**Talcott/Nvgv, 1995**) the mean priority accorded to each item was calculated and the items clustered. Clustering was limited to eight clusters, since otherwise excessively small (possibly one-item!) clusters might arise.

- The eight clusters were evenly distributed between the four quadrants formed by Cartesian X- and Y-axes. The X-axis exhibits the transition from 'feeling-oriented' to 'conduct-oriented'; and the Y-axis that from 'other-oriented' to 'self-oriented'.

- Based on the distribution of items between the eight clusters and on item priorities, the thirty 'most important' items were selected (see **Table 1**).

[3] TBS (*ter beschikking stelling*) establishments are clinics statutorily authorised to detain and treat mentally disordered offenders in The Netherlands.

- These items were set out in BSI-format (see **Table 2**), in which a more detailed and extensive description of the item content is given and five ordinal response categories are formulated. These range from complete absence of the behaviour or skill (worst-case) to its complete presence or skilled performance (best-case).

- During discussion of results it became very clear that the existing items did not give sufficient insight into the 'value' of an individual's empathic capability. Therefore for each question a threefold extra response category was added: A) credible/incredible behaviour; B) typical/untypical behaviour; C) spontaneously assessed/assessed after consultation.

- By this process an experimental research instrument consisting of 120 items was formulated.

Research sample

The empathy assessment was completed for 172 individuals, 118 of whom were assessed by two independent assessors. The control group (non-psychiatric) consisted of 51 (26 women; 25 men); mainstream psychiatric patients 24 (13 women; 10 men; 1 unrecorded); and forensic psychiatric patients 97 (6 women; 89 men; 2 unrecorded). Both psychiatric samples were considerably younger than the control group, 45 percent of whom were older than 45 years, compared with 25 percent of mainstream psychiatric patients and only 16 percent of forensic patients. The forensic group is primarily male, corresponding to the situation in The Netherlands forensic psychiatric population as a whole, where women make up less than 10 percent of the total. Later it will be seen that neither age nor gender appeared to exert a significant effect on the empathy score.

Inter-rater reliability

Inter-rater reliability for the thirty empathy questions was investigated using Kendall *tau* correlation coefficients (**N=118**). Tests showed significant agreement for all questions, with the exception of question 21 ($p=0.1518$). Twenty-three items had a significance level of $p=0.0001$. The remaining six items varied from $p=0.0002$ to $p=0.0042$. Of the associated dichotomies (A. Credible/incredible behaviour; B. Typical/untypical behaviour; C. Spontaneously assessed/assessed after consultation), twenty were deleted due to poor inter-rater reliability. Investigation using the Fisher exact test showed that, for 'credible/incredible behaviour', no significant agreement was achieved for items 3, 6, 7, 9, 15, 16, 22, 23, 24, and 28. In the case of 'typical/untypical behaviour', assessors differed in opinion on items 7, 19, 24 and 28. Finally, for 'spontaneously assessed/assessed after consultation', differences occurred for items 16, 20, 25, 26, 28 and 30. These twenty items were therefore removed from the item pool of 120 items so that 100 items remained.

Principal component analysis shows that the original thirty-dimensional concept of empathy, as illustrated by the thirty-item empathy scale, is in fact a unidimensional concept. This factor accounts for 56 percent of the variability, with which all variables are consistent. The remaining factors are unstructured "noise". Additionally the three dichotomous variables [A. reliable/unreliable; B. typical behaviour/untypical behaviour; C. spontaneously assessed/assessed after consultation] can each be described unidimensionally. The first factor [reliable/unreliable] accounts for 48 percent of the variability, counteracted by no single variable. The first factors of the dichotomies [typical behaviour/untypical behaviour] and [spontaneously assessed/assessed after consultation] respectively account for 55 percent and 52 percent of the variability. Here again there are no variables counteracting these first factors.

Age, gender and group
Because of differences in age and gender-ratio between the three research sub-samples, it was necessary to determine whether a relationship existed between age, gender and sub-sample on the one hand, and 'empathy' as operationalised in the remaining 100 items on the other. Corrected for missing observations, this check was carried out on a sample size **N=167**. When the three variables and their associated interactions were examined together using Type III sum-of-squares, only the inter-group variable appeared significant for every pairwise comparison (**p=0.0001**). Neither age (**p=0.9539**) nor gender (**p=0.4106**) showed any significant influence on empathic capability in the samples studied.

Antisocial personality disorder
There is clinical consensus that in individuals with an antisocial personality disorder empathic capability is weak or lacking. This is recognised in the DSM-IV criteria for antisocial personality disorder (APD) whose sufferers 'frequently lack empathy' (**APA, 1994**, p. 647). According to some experts differences exist between antisocial personality disorder and 'psychopathy' (see, for example, **Hildebrand and De Ruiter, 1998**). However, flaws in empathic capability feature in both descriptions.

As far as we can ascertain, no investigation has been undertaken regarding statistical support for the hypothesis that individuals with an APD classification score low on empathy. In order to determine whether such a statistical connection exists, the assessors who had collaborated in the present investigation were asked to rate individuals on criteria derived from the DSM-IV classification of antisocial personality disorder (**APA, 1994**) (see **Table 3**). All criteria were given without asking the assessor whether a classification of antisocial personality disorder actually existed in individual cases. On the scoring-sheet for the empathy scale there is a question concerning DSM-IV diagnosis(es). This information was compared with scores on the empathy scale as originally developed, in order to determine whether, within these hundred items, there was a 'hidden' factor related to the criteria for APD according to the DSM-IV. Otherwise we proceed *pace* Binet: and to the question, 'What is empathy?' comes back the simple answer: 'That which is measured by an empathy instrument'!

Relation between empathy and DSM criteria of antisocial personality disorder
Initially an unrotated factor analysis was carried out on the hundred empathy questions, in which the primary concept (empathy) became completely represented in the first factor. The results showed three variables with factor scores. All eleven DSM-variables associated with antisocial personality disorder correlate significantly with the first factor. No single DSM-variable correlates significantly with the second or third factors. These second and third factors are not discussed here.

Testing of the null hypothesis that the concepts of empathy and APD are unrelated produces a rejection. A canonical correlation between the hundred empathy variables and the eleven DSM-IV variables shows seven mutually independent, significantly correlated pairs of vectors. The methods of Wilks, Pillai, Hotelling-Lawley and Roy (see, for example, **SAS Institute, 1999**) demonstrate that empathy is strongly related to the collective elements of antisocial personality disorder. There are seven detailed propositions possible regarding latent variables hidden within these hundred-dimensional and eleven-dimensional structures, which can be explained using standardized canonical coefficients. It is unnecessary to work these out in detail within the context of the present article.

The new empathy scale

A hundred-item empathy scale is too long to act as a sixth sub-scale within the existing BSI-D (five sub-scales with a total of 120 items). Therefore a stepwise regression analysis was undertaken with the objective of deriving a shorter list by means of which the incorporated concepts of the original would be reliably represented. Such a list needed to be short; to account for sufficient variability (± 95 percent); and to avoid strong intercorrelations between its items. The stepwise regression analysis showed that, by step six, 94 percent of the variability was explained, and all four aspects of the original list were represented. Of the six items four are so-called 'letter-items' (as previously indicated, the list of 100 variables consists of 29 empathy items and 71 letter-items, the latter including: A. reliable/unreliable; B. typical behaviour/untypical behaviour; C. spontaneously assessed/assessed after consultation). It is necessary to make a complete hierarchy from the list because otherwise questions regarding the contribution of letter-items cannot be answered. By making the new list completely hierarchal - that is to say, by appending the letter-items to the empathy items - it was possible to develop a ten-item list explaining 95 percent of the variability (**Table 4**).

Uncorrected, the new scale offers no useful discrimination between the groups, and is still some way from an operational empathy concept. The average empathic capability of the total group investigated offers no usable criterion of empathic capability, because the group of forensic patients constitutes almost 60 percent of assessed individuals; and the scores for this group on the empathy scale are significantly lower than those of the control group. Therefore we are some way from calculating a 'normal' empathic capability, since logically this would be the mean for the control group.

There remains the problem that not all items are scored identically: the empathy items are scored on a five-point ordinal scale, whilst the letter-items are scored dichotomously. The scalar scores should be calculated without adaptation. Then calculating in integers, starting from the middle (zero point) of the control group, the empathy scale is adapted as follows:

$$\textbf{Items}\ (1 + 2 + 3 + 4 + 5 + 6) - 4(3A + 4B + 5B + 6C) - 6$$

[**NOTE:** The variables [1,2,3,4,5,6] are renumbered from the original scalar items [1,16, 29,19,26, 4]

First the sum of scores for the six empathy items is calculated, giving the uncorrected score of empathic capability. From this sum a correction factor (four times the sum of the four letter-items) is subtracted, because this represents the real antithetical loading on the factor. Finally from the result of this calculation an additional correction factor of 6 is subtracted, derived from the mean for the control group.

This gives a skewed distribution (see **Figure 1**), clearly showing that empathic capability of forensic psychiatric patients is lower than that of mainstream psychiatric patients; and still lower than that of a control group of non-psychiatric patients.

Synopsis

Concept mapping has shown itself to be a practical, useful procedure for quickly assembling a comprehensive item pool relating to 'empathy'; and for developing a prototype empathy scale for preliminary investigation. In order to obtain a reliable focus on the empathic capability of forensic psychiatric patients, a correction for 'credibility' of observed behaviours must be applied. For purposes of forensic psychiatric treatment it is also important to gain insight

into the degree of independence, spontaneity or help required by an individual in producing empathic behaviour.

The canonical correlation between the hundred-item empathy scale and the eleven DSM-IV variables supplies seven mutually independent, significantly correlated pairs of vectors, explicable in terms of the standardized canonical coefficients.

If repeated measurement is desired, a practical, applicable empathy scale must contain the minimum possible number of items. Using techniques of multivariate analysis, it eventually became possible to derive a ten-item empathy observation instrument retaining 95 percent reliability of the original appraisal. Similarly, scalar scoring has been modified so that it can be completed very easily in a paper-and-pencil version.

Discussion

Systematic, reliable and objective practical assessment of the empathic capability of forensic psychiatric patients using an observational instrument may be expected to bring clear added value. It provides a useful addition to the subjective clinical judgment of the assessor or clinical team, which frequently appears to correlate only moderately with actual recidivism rates (**Hanson and Bussière, 1998**). At present such subjective judgment is practically always the only measure of empathic capability and remains necessarily implicit. By contrast, assistance provided by the proposed empathy scale is explicit, making even less obvious comparisons systematically possible. Naturally at this early stage, the precise value of the instrument as a predictor of risk or recidivism is not yet finally determined.

Since analogous observations are recorded by each assessor, repeated measurement eventually supplies an 'image' of observable changes. This in turn highlights any desired or unwanted effects of the treatment being offered and supplies useful indications for continuing treatment and/or the necessity for specific case-handling. Finally in the longer term, it becomes obvious what relationships exist between treatments, observed empathic behaviours and recidivism rates in individuals and groups. The present study has developed a useful instrument with demonstrably good inter-rater reliability. As a practical instrument it must now be further tested in terms of its clinical, predictive and methodological usefulness.

REFERENCES:

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders (4th ed.).* Washington: American Psychiatric Association.

Barratt-Lennard, G.T. (1962). Dimensions of therapist response as causal factors in therapeutic change. *Psychological Monographs*: *General and Applied*, 76, 43, 562, 1-33.

Binder, U. (1999). *Empathie en empathieontwikkeling met betrekking tot vroege, ernstige stoornissen.* (Empathy and empathy development with relationship to early serious disturbances). *Tijdschrift Cliëntgerichte Psychotherapie, 37*, 266-278.

Boer N. de (1997). *Verkennen begint met een betrouwbare kaart*. (Exploration begins with a reliable map). *Tijdschrift voor de Sociale Sector, 5*, 14-17.

Boevink, W. & Wolf, J. (1998). *Om de kwaliteit van het bestaan. Betekenis en toepassing van het begrip 'kwaliteit van leven' in de psychiatrie.* (On the quality of being: meaning and application of the concept "quality of life" in psychiatry). *De Psycholoog*, July/August, 308-313.

Brody, S. (1995). Patients misrepresenting their risk factors for AIDS. *International Journal for STD-AIDS*, 6, 392-398.

Bühler, K.E., Haltenhof, H., Geyer, S., Bardeleben, H. (1999). Reliability of biographical data, their relations to personality variables and their influence on life events. *Journal of Affective Disorders*, 53 (1), 67-76.

Burns, D.D. & Auerbach, A. (1996). Therapeutic empathy in cognitive-behavioral therapy: does it really make a difference? In: **Saedrooskin, P.M.**, *Frontiers of Cognitive Therapy*. New York, The Guilford Press.

Coumans, G.G.H., Wilde, G. de & Hull, H.R. (1994). Concept Mapping. Kwaliteit in kaart gebracht (Concept Mapping: bringing quality onto the map). , GGD-Nieuws, 12-15.

Cronbach, L.J. (1968). *Essentials of Psychological Testing*. New York: McGraw-Hill.

Derks, F. & Mulder, J. (1993). *Het ontwikkelen van criteria voor de evaluatie van een forensische deeltijd- en polikliniek.* (Development of criteria for evaluating a forensic part-time and polyclinic). *Tijdschrift voor Psychiatrie, 35*, 208-214.

Dienst Justitiële Inrichtingen **(Department of Judicial Planning) (1998).** *"Een goed aanbod voor TBS-verpleegden die langdurig in de TBS verblijven kan niet zonder ...".* ("Adequate long-term accommodation for TBS patients cannot be guaranteed unless ... " *Directie beleidszaken* (Directorate of Business Policy), December, 1998.

Duan, C. & Hill, C.E. (1996). The current state of empathy research. *Journal of Counseling Psychology*, 43, 3, 261-274.

Eslinger, P.J. (1998). Neurological and neuropsychological bases of empathy. *European Neurology*, 39, 193-199.

Gould, D. (1990). Empathy: a review of the literature with suggestions for an alternative research strategy. *Journal of Advanced Nursing, 15*, 1167-1174.

Hanson, R.K. & Bussière, M.T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology*, 66(2), 348-362.

Hart, W. & Tomazic, T. (1999). Comparison of percentile distributions for anthropometric measures between three data sets. *Paper presented at Annual Joint Statistical Meeting*, Baltimore, MD.

Hildebrand, M. & Ruiter, C. de (1998). *Ontwikkelingen in het onderzoek naar psychopathie. Psychodiagnostisch gereedschap*. (Research developments in psychopathy: psychodiagnostic instruments). *De Psycholoog*, July/August, 314-320.

Hogan, R. (1969). Development of an empathic scale. *Journal of Consulting Clinical Psychology*, 33, 307-316.

Ickes, W. (1997). Introduction. In **Ickes, W.** Empathic Accuracy. New York/London: The Guilford Press.

La Monica, E.L. (1980a). *Empathieprofiel* (The La Monica Empathy Profile). Velp/Mechelen: TFC TrainingsMedia.

La Monica, E.L. (1981). Construct validity of an empathy instrument. *Research in Nursing and Health, 4*, 389-400.

La Monica, E.L. (1994). The La Monica Empathy Profile. New York: Xicom/Tuxedo.

Lewontin, R.C. (1995). Sex, Lies and Social Science. New York: Rev Books, 42, 24-29.

Nijssen, Y.A.M., Schene, A.H. & Haan R.J. de (1999). *Kwaliteitsbeoordelingsinstrumenten voor patiënten in de geestelijke gezondheidszorg. Een overzicht.* (Quality assessment instruments for patients in psychiatric health care: an overview.). *Tijdschrift voor Psychiatrie, 41*, 11, 651-663.

Reynolds, W.J. & Presley, A.S. (1988). A study of empathy in student nurses. *Nurse Education Today*, 8, 123-130.

SAS Institute (1999): The SAS/STAT User's Guide (Multivariate Analysis of Variance). Cary, NC: SAS Institute.

Schachter, D.L. (1999). The seven sins of memory: insights from psychology and cognitive neuroscience. *American Psychology*, 54, 182-203.

Spil, F. (1999). *De betrouwbaarheid, validiteit en bruikbaarheid van een nieuwe empathie beoordelingsschaal.* (Reliability, validity and utility of a new empathy assessment scale.). *Doktoraal skriptie Geestelijke Gezondheidskunde* no. 972796 (Doctoral dissertation in psychiatric health care studies no. 972796). University of Maastricht.

Strien, A.P.M. van (1999). *Empathie. Een kwaliteitsaspect van de psychoanalytische praktijk? Een exploratief onderzoek van een concept.* (Empathy: a qualitative aspect of psychoanalytic practice? Exploratory investigation of a concept.). Amsterdam: Thela Thesis Publications.

Talcott bv/NcGv (1995). *Handboek Concept Mapping met Ariadne*. (Manual of Concept Mapping using Ariadne). Linde: Paardelanden 18.

Trochim, W.M.K. (1989). An introduction to concept mapping for planning and evaluation. *Evaluation and Program Planning, 12*, 1-16.

Vanaerschot, G. (1999). *De empathische interactie in de praktijk.* (Empathic interaction in practice). *Tijdschrift Cliëntgerichte Psychotherapie, 37*, 5-20.

Van Dale (1999): *Groot Woordenboek* (Great Dictionary of the Dutch Language). Utrecht/Antwerp: Van Dale Lexicografie.

Van Erven, A.C.J.M. (1999). *Behavioural Status Index - Nederlandse versie (BSI-D).* Nederlandse vertaling en bewerking van **Reed, V., Woods, P. and Robinson, D. (2000):** The Behavioural Status Index (BSI) – a 'life skills' assessment for selecting and monitoring therapy in mental health care. United Kingdom: Psychometric Press. (Dutch translation and adaptation of the English text).

Wiseman, T. (1996). A concept analysis of empathy. *Journal of Advanced Nursing, 23*, 1162-1167.

Dr A.C.J.M. van Erven is a clinical psychologist, psychotherapist and researcher working in the Forensic Psychiatric Circuit and the Circuit Management of GGzE (Geestelijke Gezondheidszorg Eindhoven en de Kampen), Eindhoven, The Netherlands.

Dr Jan B. Dijkstra is Head of the Information and Communication Technology Group for Education and Research, The Technical University, Eindhoven, The Netherlands.

Prof Dr Val Reed is visiting professor in forensic mental health, School of Health and Social Care, Sheffield Hallam University, United Kingdom.

Table I: The original (thirty-item) empathy scale.

Item:	Descriptor:
1	Imagining oneself in the life-world of another person
2	Understanding the feelings of another person, distinct from those of oneself
3	Sensitivity to others
4	Expressing sympathy for the wishes and needs of another person
5	Pleased for others
6	Allows others to express themselves
7	Interested in social 'give-and-take'
8	Dealing with conflict
9	Sharing conversations
10	Curbing self-interest
11	Listening to others
12	Physical 'mirror responses'
13	Offering support
14	Avoiding abuse
15	Listening and questioning
16	Expressing regret
17	Accepting ideas
18	Comforting others
19	Acknowledging the victim
20	Giving others 'breathing space'
21	The victim as a person
22	Concern for others' troubles
23	Psychological 'intrusion'
24	Sharing 'terrors'
25	Expressing consideration
26	Taking an interest
27	Asking about feelings
28	Making eye contact
29	Balancing interests
30	Doing things for others

Table 2: Items 1 to 3 of original empathy scale as used prior to current study

1. IMAGINING ONESELF IN THE LIFE-WORLD OF ANOTHER:

1.1	1.2	1.3	1.4	1.5
Never shows evidence of imagining another's life-world	**Seldom** shows evidence of imagining another's life-world	Shows **average** evidence of imagining another's life-world	**Frequently** shows evidence of imagining another's life-world	**Always** shows evidence of imagining another's life-world

The assessor must observe a clear, typical behaviour indicating that the patient is able to identify him/herself with the feelings and thoughts of another person.

Life-world means the conscious experience, thoughts and feelings; in this case, of a person other than the patient.

Never: Never observed.
Seldom: Observed only infrequently.
Average: Observed as often as not (50:50).
Frequently: Observed more often than not.
Always: Observed on every occasion.

---

2. UNDERSTANDING THE FEELINGS OF ANOTHER PERSON, DISTINCT FROM THOSE OF ONESELF.

2.1	2.2	2.3	2.4	2.5
Never shows evidence of understanding another's feelings	**Seldom** shows evidence of understanding another's feelings	Shows **average** evidence of understanding another's feelings	**Frequently** shows evidence of understanding another's feelings	**Always** shows evidence of understanding another's feelings

The assessor must observe a clear, typical behaviour indicating that the patient understands the feelings of another person; and can view these separately from his/her own feelings.

Understands the feelings of another person: This is the ability at will to understand the emotional reactions of another person; and (at least partly) to accept these and conform one's behaviour to them.

Never: Never observed.
Seldom: Observed only infrequently.
Average: Observed as often as not (50:50).
Frequently: Observed more often than not.
Always: Observed on every occasion.

---

3. SENSITIVITY TO OTHERS:

3.1	3.2	3.3	3.4	3.5
Never shows evidence of sensitivity to others	**Seldom** shows evidence of sensitivity to others	Shows **average** evidence of sensitivity to others	**Frequently** shows evidence of sensitivity to others	**Always** shows evidence of sensitivity to others

The assessor must observe a clear, typical behaviour indicating that the patient can be sensitive to another person's feelings. The patient shows signs of fellow-feeling: for example, by his/her style of seeking information; by making supportive remarks; or by sharing a joke with the other person.

Never: Never observed.
Seldom: Observed only infrequently.
Average: Observed as often as not (50:50).
Frequently: Observed more often than not.
Always: Observed on every occasion

Table 3: Criteria of classification for antisocial personality disorder (APD) based on the DSM-IV[4].

Criterion:	Descriptor:
A1	There is a pervasive pattern of disregard for and violation of the rights of others.
B2	This pattern has been present from the age of fifteen years.
3	Failure to conform to social norms with respect to lawful behaviors as indicated by repeatedly performing acts that are grounds for arrest.
4	Deceitfulness, as indicated by repeated lying, use of aliases, or conning others for personal profit or pleasure.
5	Impulsivity or failure to plan ahead.
6	Irritability and aggressiveness, as indicated by repeated physical fights or assaults.
7	Reckless disregard for safety of self or others (for example, by driving under the influence of alcohol or drugs).
8	Consistent irresponsibility, as indicated by repeated failure to sustain consistent work behavior or honour financial obligations.
9	Lack of remorse, as indicated by being indifferent to or rationalizing having hurt, mistreated, or stolen from another.
C10	The individual is at least age 18 years, and there is evidence of conduct disorder with onset at or before age 15 years.
D11	The occurrence of antisocial behavior is not exclusively during the course of schizophrenia or a manic episode.

(After DSM-IV: In order for the classification of Antisocial Personality Disorder (APD) to be applied, there must be (A) a pervasive pattern of behaviour (point 1); (B) recognisable from the age of fifteen years (point 2); involving three or more of the behaviours described in points 3 to 9, ***together with (C) and (D)*** (points 10 and 11).

[4] **American Psychiatric Association (1994).** *Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition).* Washington: American Psychiatric Association.

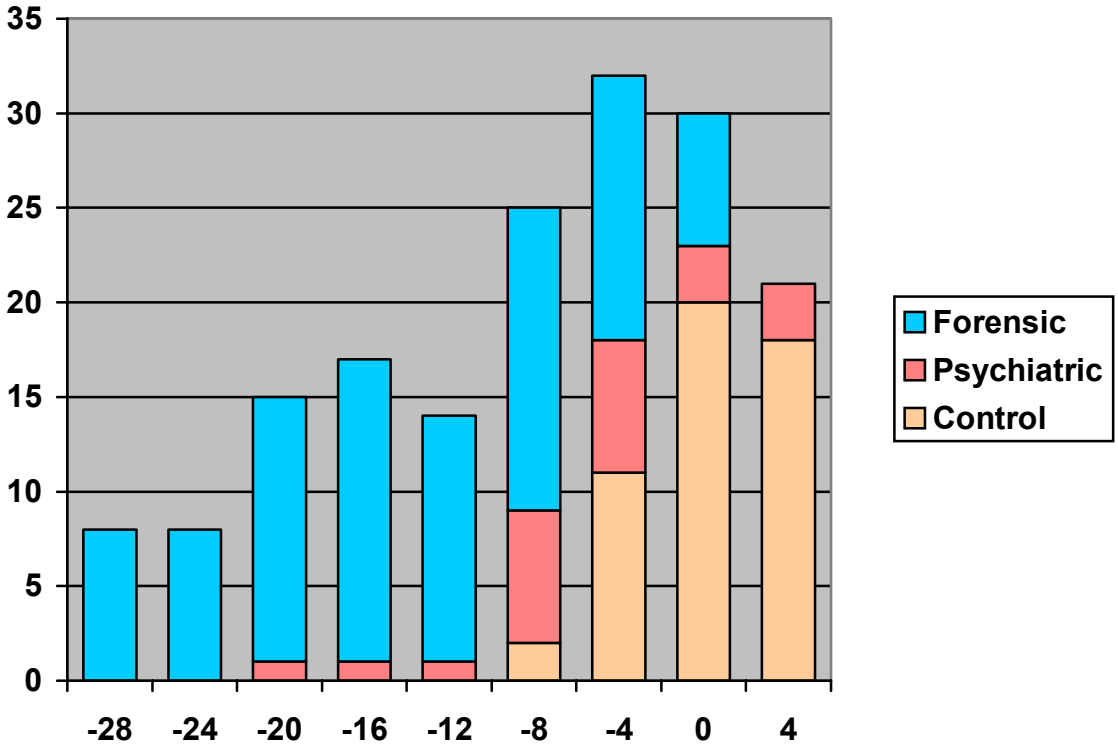
Table 4: The revised (ten-item) empathy scale.

The patients shows evidence of:

1	**Imagining him/herself in the life-world of another person.**				
	1 Never	2 Seldom	3 Average	4 Frequently	5 Always
2	**Expressing regret.**				
	1 Never	2 Seldom	3 Average	4 Frequently	5 Always
3	**Balancing interests.**				
	1 Never	2 Seldom	3 Average	4 Frequently	5 Always
3A	***if Item 3 score is between 2 and 5 then***				
	1) credible behaviour		2) not credible behaviour / only done for personal advantage		
4	**Recognising the victim as a person.**				
	1 Never	2 Seldom	3 Average	4 Frequently	5 Always
4B	***if Item 4 score is 1 then***				
	1) typical behaviour		2) untypical behaviour		
5	**Taking an interest in others.**				
	1 Never	2 Seldom	3 Average	4 Frequently	5 Always
5b	***If Item 5 score is 1 then***				
	1) typical behaviour		2) untypical behaviour		
6	**Expressing sympathy for the wishes and needs of another person.**				
	1 Never	2 Seldom	3 Average	4 Frequently	5 Always
6C	***if Item 6 score is between 2 and 5 then***				
	1) spontaneously assessed		2) assessed after consultation		

Variables [1,2,3,4,5,6] are renumbered from original numbers [1,16,29,19,26, 4]

Figure 1: Distribution of corrected scores on the ten-item empathy scale.



Control group (N=51); Psychiatric patients (N=24); Forensic psychiatric patients (N=97)