The Minnesota Multiphasic Personality Inventory - 2 in a Dutch pre-trial forensic population:

Personality profiles and gender comparisons

Master's thesis

Joris de Bruijn

Anr.: 168339


Thesis supervisor: dr. J. J. Sijtsema

Second assessor: prof. dr. S. Bogaerts

Tilburg University

Tilburg School of Social and Behavioral Sciences

Department of Developmental Psychology


Research supervisor: dr. M. Kempes

Netherlands Institute for Forensic Psychiatry and Psychology

## Abstract

Previous research on Minnesota Multiphasic Personality Inventory - 2 (MMPI-2) personality profiles in criminal offender populations yielded mixed results regarding the number of different profiles and the depth of information provided by these profiles. This study aimed to distinguish meaningful personality profiles based on the MMPI-2 in a Dutch forensic pre-trial population. Cluster analysis was performed on the MMPI-2 questionnaires completed by 1663 defendants (249 female, 1414 male) awaiting trial, in the period of 2006-2014. The preferred cluster solution for the total population consisted of three clusters, which included a non-disturbed, mildly disturbed and very disturbed clusters. Viewed separately for the female and male subpopulation, 2-clusters provided the most useful clinical profiles: a non-disturbed and (very) disturbed cluster for both genders. This study also discussed an additional 4-cluster solution and gender comparisons for the 2-, 3-, and 4-cluster solutions. The female subpopulation scored higher on Clinical Scales 6, 7, and 8; whereas the male subpopulation scored higher on Clinical Scales 4 and 5. The author compared the results with a similar previous study and discussed conclusions regarding future use of the MMPI-2 as an appropriate personality assessment. Additional research into the fit of the profiles with possible Axis I and Axis II diagnoses is recommended.

Keywords: MMPI-2, personality profiles, classification, forensic population, offender population, cluster analysis.

## Samenvatting

Voorgaande studies naar persoonlijkheidsprofielen gebaseerd op de Minnesota Multiphasic Personality Inventory -2 (MMPI-2) varieerden sterk in het aantal verschillende profielen en de bruikbaarheid van de informatie van de profielen. Dit onderzoek had als doel om verschillende, betekenisvolle persoonlijkheidsprofielen te onderscheiden, in een Nederlandse populatie van forensisch onderzochte verdachten. Door middel van cluster analyse werden MMPI-2 vragenlijsten geanalyseerd die door 1663 verdachten (249 vrouwen, 1414 mannen) in de periode 2006-2014 waren ingevuld. Voor de totale populatie bleken drie clusters het meest geschikt, bestaande uit een niet-gestoord, licht gestoord en ernstig gestoord cluster. Voor de vrouwelijke en mannelijke populatie apart bleken twee clusters de meest betekenisvolle te verschaffen over persoonlijkheidsprofielen: een niet-gestoord en een (ernstig) gestoord cluster voor beide geslachten. Dit onderzoek besprak ook een aanvullende 4-cluster oplossing, evenals vergelijkingen tussen de vrouwelijke en mannelijke subpopulaties op de 2-, 3- en 4-cluster oplossingen. De vrouwelijke subpopulatie scoorde hoger op Klinische Schalen 6, 7 en 8. De mannelijke subpopulatie scoorde hoger op Klinische Schalen 4 en 5. De auteur vergelijkt de huidige resultaten met een eerdere soortgelijke studie en bespreekt conclusies ten aanzien van toekomstig gebruik van de MMPI-2 als persoonlijkheidsvragenlijst. Toekomstig onderzoek naar de koppeling met eventuele As I en As II diagnoses is aanbevolen.

Trefwoorden: MMPI-2, persoonlijkheidsprofielen, classificatie, forensische populatie, delinquente populatie, cluster analyse.

**Introduction**

A considerable part of both pre-trial and detained forensic populations suffers from a variety of personality disorders. In an exhaustive literature review of studies performed in twelve countries, Fazel and Danesh (2002) found that 65% of the male prisoner population and 42% of the female prisoner population suffered from one or more personality disorders. As the presence of a personality disorder can have implications for the degree of accountability for a committed crime and, consequently, for the type and severity of the legal penalty and/or mental health care programs, reliable and effective personality assessments are a necessity for forensic psychological research. A widely used instrument for the assessment of personality and personality disorders in pre-trial forensic populations is the Minnesota Multiphasic Personality Inventory-2 (MMPI-2) (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989). The present study aimed to determine whether the MMPI-2 is useful in a pre-trial forensic population regarding the ability to distinguish various personality profiles. Moreover, this study focused on possible gender differences in MMPI-2 personality profiles.

The original version of the Minnesota Multiphasic Personality Inventory (MMPI) was aimed at constructing a valid measure of personality structure and psychopathology in adults (e.g., Hathaway & McKinley, 1940; Hathaway & McKinley, 1942). The MMPI served as an inventory that was not based on or driven by any prevailing personality theory at that time. Instead it was based on empirical evidence derived from common diagnoses in clinical settings. The MMPI-2 (Butcher et al., 1989), which was the focus of the present study, has been modified to incorporate the most up-to-date research findings (e.g., Butcher et al., 2001) and currently consists of 574 items covering ten Clinical Scales (see Method section).

The use of self-report measures, such as the MMPI-2, in forensic settings has been the

subject of debate in a number of research areas. On the one hand, several studies validated the

MMPI or MMPI-2 measures in forensic settings by discerning various personality profiles (e.g.,

Espelage et al., 2003; Hall, Graham, & Shepherd, 1991; Spaans, et al., 2009). For example,

Nieberding et al. (2003) found seven different personality profiles. Apart from three profiles with

non-clinical levels, Nieberding et al. (2003) discerned one Psychopathic Deviate profile, one

profile constituting elevated scores on Psychopathic Deviate, Paranoia, Psychasthenia, and

Schizophrenia traits, and two profiles constituting elevated scores on Hypochondriasis,

Depression, and Hysteria. Another study identified ten distinct personality profiles (Megargee,

Bonn, Meyer, & Sink, 1979; Megargee, Carbonell, Bohn, & Sliger, 2001), including amongst

others the 1-2-3 profile code (Hypochondriasis, Depression and Hysteria) which reflects the so-

called 'Neurotic Triad', also found in the Nieberding et al.'s (2003) study. Based on their study,

Megargee et al. (1979) even derived a widely recognized MMPI-2 classification system for

incarcerated populations, indicating legitimized use of the MMPI-2 with imprisoned offenders.

However, other studies did not find such explicit results. In contrast to Nieberding et al.

(2003) and Megargee et al. (1979, 2001), several studies found only few distinct profiles and

hence raised doubts regarding the effectiveness of the MMPI-2. For example, one study only

found an impulsive and antisocial cluster and a cluster with more psychiatric disturbance in a

group of sexual offenders (Hall et al., 1991); and two separate studies found one (non-disturbed)

cluster with no psychological distress versus one (disturbed) cluster with varying forms of

psychopathology (Espelage et al., 2003; Spaans et al., 2009). The MMPI-2, originally, was not

specifically developed for use in forensic settings (Weiner, Freedheim & Graham, 2003), which

has raised doubts concerning its applicability in forensic populations. Other studies have

identified more shortcomings of the MMPI-2. First, the considerable length of the MMPI-2

requires the respondent to maintain focus and attention for quite a considerable amount of time. Considering, for example, the relatively high prevalence of ADHD symptoms in forensic settings, e.g., 40% of adult male longer-term prison inmates (Ginsberg, Hirvikoski & Lindefors, 2010) and 50% of female prisoners (Edvinsson, Bingefors, Lindström & Lewander, 2010), it is expected that attention diminishes during the administration of the test more profoundly than in general population samples. Second, pre-trial forensic populations have a more pronounced tendency to either exaggerate and/or feign certain symptoms such as crime-related amnesia, which is known as malingering (Cima & Van Oorsouw, 2013; Van Oorsouw & Merckelbach, 2011); or to fake good qualities, sometimes even to the point of 'supernormality' which occurs when a "patient […] might try to deny […] experiences in an attempt to make a healthy impression on his evaluators" (Cima et al., 2003: 236). Self-report measures, by nature, facilitate these response distortions.

Grover (2011) summarized pros and cons regarding the use of the MMPI-2 in correctional facilities. Administration of the test can be done efficiently and in a standardized manner. Combined with the widespread use in general populations, the long history of research validating the MMPI-2, and previous success obtained in identifying those that are mentally ill or emotionally disturbed, it is appealing for penitentiaries and correctional health facilities to administer the test. However, MMPI-2 based classification systems such as the one developed by Megargee et al. (1979, 2001) are not well suited for predicting violent behavior (Heilbrun & Heilbrun, 1995). They can at best only distinguish the more hostile offenders from less hostile offenders (Grover, 2011). In sum, these contrasting views regarding the use of the MMPI-2 fuel the need for further research into its specific application for forensic populations.

Personality structures and the MMPI-2 in forensic populations have been investigated

across different countries (e.g. Fazel & Danesh, 2002), in different ethnic groups (e.g.

McGilloway, Hall, Lee, & Bhui, 2010), and in specific types of offenders such as sexual

offenders (e.g. Davis & Archer, 2010). Large scale research focusing on gender differences in

forensic populations with regard to personality structures can be regarded as incomplete. It is

well-established that the majority of the incarcerated population consists of male offenders. For

example, in 2012 in The Netherlands, male offenders accounted for 92.4% of the incarcerated

population versus 7.6% female offenders (Dienst Justitiële Inrichtingen, 2013, p. 20).  Female

offenders in the U.S. account for less than 5% of the incarcerated population (Megargee, Mercer,

& Carbonell, 1999). The research literature shows mixed results regarding increase or decrease

of the incarcerated population, and the relative increase or decrease in male versus female

offenders. For example, in the period of 1973 to 1993, the incarcerated population in the U.S.

grew by 446%. The rate of increase among female offenders was almost twice as high as the rate

among male offenders in the U.S.: 846% versus 435%, respectively (Maguire & Pastore, 1995).

However, in The Netherlands, the total number of incarcerated offenders decreased, by 7.1% in

the period of 2008 to 2012. Additionally, the inflow of female offenders decreased by 16%; the

inflow of the incarcerated male population decreased by 6% (Dienst Justitiële Inrichtingen, 2013,

p. 20).

The lower number of female prisoners compared to male prisoners notwithstanding, the

female incarcerated population deserves research for salient personality features as well. As

mentioned earlier, personality disorders are highly prevalent in this population (Fazel & Danesh,

2002). Attempts to classify personality structures in female offender populations have been

based on systems that were developed for male offenders (Megargee, 1997; Megargee et al.,

1999). This led to some surprising results. The Masculinity/femininity scale of the MMPI-2

revealed that both men and women scored in the masculine direction (Megargee et al., 1999).

Both genders showed clinical elevations on the Psychopathic deviate, Paranoia, and Hypomania

scales. Furthermore, female offender MMPI-2 scores were significantly higher than those of their

male counterparts. Megargee et al. (1999) concluded that incarcerated women seem more

socially deviant than incarcerated men. Other studies attempted to shed light on the female

imprisoned population as well. For example, adult women who were charged with either murder

of their child, murder of their husband or murder of an unrelated adult did not differ significantly

from each other on all the clinical scales of the MMPI-2 (McKee, Shea, Mogy, & Holden, 2001).

Also, women, more often than men, are found not guilty by reason of insanity (Hatters Friedman,

Hall, & Sorrentino, 2013). This makes women who have committed a homicide, for example, a

very interesting and useful focus of study in terms of personality, motivation for violent acts, and

so on: the judicial perspective on homicidal women differs from the perspective on homicidal

men. Theories regarding higher prevalence rates of disorders in women in general have been

both defended and debated. Several studies have found higher prevalence rates for depressive

and anxiety disorders in women, whereas an antisocial personality and substance use disorders

would be more prevalent in men. Still, mental disorders in general would have equivalent

lifetime prevalence rates (e.g., Karno et al., 1987, Robins et al., 1984).  Aneshenzel, Rutter and

Lachenbruch (1991) emphasized that the impact of stress, rather than specific disorders, can

render individuals vulnerable to contact with the justice system. Women usually endure more

social and affective stress during the life span, which has its effect on behavior and intensity of

interpersonal style. Supposedly, these relatively higher amounts of experienced stress in women

would be contributing mostly to contact with the justice system, rather than personality disorders

alone. However, stress would still evoke more significant problems with the justice system in

women with personality disorders than in women without a personality disorder (Aneshenzel et al., 1991).

In light of the mixed results of previous research regarding the ability of the MMPI-2 to differentiate between distinct personality profiles in forensic populations and the gap in the literature regarding MMPI-2 outcomes in female offenders, the aim of the present study was to explore the suitability of the MMPI-2 in a Dutch pre-trial forensic population. Spaans et al. (2009) examined MMPI-2 outcomes in a Dutch pre-trial forensic male population committed to the Pieter Baan Centrum, where defendants are placed under clinical observation and research in anticipation of their trial. Two clusters were found as the most preferable outcome. One cluster indicated severe psychological distress, and the other cluster reported no personal/emotional traits and symptoms. This study aimed to replicate, yet also extend Spaans et al.'s (2009) study by 1) investigating both male and female defendants, that 2) received an ambulant psychological evaluation; ambulant, meaning that the defendants were not in a controlled research environment (such as the Pieter Baan Centrum) during their psychological evaluation. It can be argued that defendants awaiting trial do not show their true colors, as there is still a verdict to be formulated by the judge. The main research question was whether the MMPI-2 can differentiate between various personality profiles in a pre-trial forensic population. Based on Spaans et al. (2009), the first research hypothesized was that there will be one cluster indicating psychological distress and one cluster indicating the absence of psychological distress. The second hypothesis was that the cluster with psychological distress would have clinically elevated scores on the Psychopathic deviate, Paranoia, and Schizophrenia scales (Megargee et al., 1979, 2001; Nieberding et al., 2003). The second aim of this study was to examine the role of gender on MMPI-2 personality profiles. There were no specific expectations regarding any possible differences in personality

profiles on Clinical Scale level; following Megargee et al.'s (1999) reasoning, however, the third hypothesis was that, overall, women would obtain significantly higher scale scores than men.

## Method

### Procedure

*Sampling procedure and sample size*. All participants in this study had already completed the MMPI-2 - Dutch adaptation (Sloore, Derksen, De Mey, & Hellenbosch, 1996) in light of their pre-trial psychological evaluation. The MMPI-2 data used in the present study were registered MMPI-2 scores in computers at various sites of the Netherlands Institute for Forensic Psychiatry and Psychology (NIFP) from January 2006 - June 2014. Assessment of the MMPI-2 occurred either in a House of Detention, at an NIFP location, or in the defendant's home. Unfortunately, due to privacy concerns, this study did not have access to information regarding the specific assessment location, the (absence of a) diagnosis as the end result of the psychological evaluation, index crime, level of education, and socio-economic status, per defendant.

There is no consensus in the research literature regarding appropriate sample sizes when employing cluster analysis. As a rule of thumb, Formann (1984) recommended a sample size of $2^m$, where *m* equals the number of clustering variables. This study included ten clustering variables: the ten Clinical Scales of the MMPI-2 (see *Measurements*). Following Formann's (1984) rule of thumb, this amounts to a minimum sample size of $2^{10}$: 1024 participants.

*Participants*. All participants in this study were defendants awaiting trial for the index crime they had been charged with. The original sample contained the MMPI-2 profiles of 1782 participants. As this study served to investigate MMPI-2 profiles in the general pre-trial forensic population, participants were included simply for having completed the MMPI-2 during their

pre-trial psychological evaluation. Eventually, 1663 participants were included, which exceeds

the required minimum sample size of 1024 of Formann's (1984) rule of thumb. Criteria for

excluding participants were missing information concerning gender and incomplete MMPI-2

profiles (when 35 or more items had not been answered by the participant). Some participants

had been assigned a 'random age' (e.g. an age of 112 years old) to preserve privacy; these

participants were still included in this study as age was not of primary interest. The total

population ($N$ = 1663) had a mean age of 34.3 years (range: 18-84 years, $SD$ = 13.4); the

aforementioned outliers were excluded from these descriptive statistics. The female

subpopulation (15.0%, $n$ = 249) had a mean age of 34.5 years (range: 18-84 years, $SD$ = 12.6),

the male subpopulation (85.0%, $n$ = 1414) had a mean age of 34.3 years (range=18-82 years, $SD$

= 13.5).

    *Measurements.* The MMPI-2 (Butcher et al., 1989) is a self-report instrument that

consists of 574 items and includes statements that participants rate as either true (1) or false (0).

The ten Clinical Scales assessed are (1) Hypochondriasis (Hs, 32 items) that focuses on concern

and preoccupation with health and bodily symptoms; (2) Depression (D, 57 items), focusing on

depressive symptoms; (3) Hysteria (Hy, 60 items), focusing on emotionality and perceived

personal problems and vulnerability; (4) Psychopathic Deviate (Pd, 50 items), measuring e.g.

need for control, adherence to society's rules, anger, and struggle and conflict; (5)

Masculinity/Femininity (Mf, 56 items), concerning stereotypical masculine or feminine interests

and behaviors; (6) Paranoia (Pa, 40 items), focusing on a person's (in)ability to trust others,

sensitivity, and suspiciousness; (7) Psychasthenia (Pt, 48 items), which focuses on a person's

level of experienced tension and anxiety, worry, and obsessiveness; (8) Schizophrenia (Sc, 78

items), concerning social alienation, and unusual/odd cognitive, perceptual, and emotional

experiences; (9) Hypomania (Ma, 46 items), measuring a person's level of energy and excitability; and (0) Social Introversion (Si, 69 items), focusing on the level of (dis)comfort of being around other people. Scales 1, 2, 7, and 8 are known as symptom scales, whereas scales 3, 4, 5, 6, 9 and 0 are known as character scales reflecting a person's traits. The raw scores were transformed to *T*-scores. Per scale, a classification of *T*-scores indicates the severity of the aforementioned symptoms. *T*-scores have a mean of 50 and a standard deviation of 10. MMPI-2 *T*-scores above 64 are generally considered to be of clinical relevance (Sloore et al., 1996).

To assess the validity of the scale profiles, the MMPI-2 has several validity scales indicating whether a participant was, for example, over- or underreporting complaints or symptoms, lying, or answering in a persistent pattern (e.g. marking ten consecutive items "false" without paying attention to the actual item content). The MMPI-2 includes a Lie (L), Infrequency (F), and Correction (K) scale for each profile (see Table 1). The L scale concerns naive attempts to place oneself in a morally or culturally favorable light by denying moral imperfections; it reflects a naive form of defensiveness and impression management, with a reluctance to admit common shortcomings. The F scale can reveal the tendency to claim highly unusual attitudes and behaviors as a function of severe psychopathology; it concerns rare responses, which serve to place oneself in an unfavorable light. The K scale assesses the tendency to control and limit disclosure of distress, discomfort, and problems relating to others, denying problems, and reluctance to admit symptoms. Again, *T*-scores above 64 are considered to be of relevance for assessing invalidity (and for the L and K scales, *T*-scores below 40). Additionally (and apart from whether the *T*-scores are elevated), the way these three scales relate to one another can also reveal important information. A low score on L, a high score on F, and a low score on K (shaped graphically as an upward arrow) indicates participants' awareness of their own problems. A high

score on L, a low score on F, and a high score on K (shaped graphically as a downward arrow) reveals possible denial or underreporting of complaints/symptoms, or presenting oneself in a favorable manner.

When doubts of profile validity are raised, the True Response Inconsistency (TRIN), Variable Response Inconsistency (VRIN), F-back (Fb), Infrequency-Psychopathology (Fp), and the Superlative self-presentation (S) scores can be consulted (see Table 1). TRIN measures the tendency to answer all or most items with true. VRIN measures carelessness and random responding. Fb is the F-scale (see above) for items 371-574. Fp measures faking bad and random responding. S gives an indication of defensive responding, underreporting, and the extent to which participants present themselves in an overly positive manner. No single validity check can be used on its own to determine whether a profile is to be rendered invalid. All validity checks must be taken into consideration (Sloore et al., 1996).

*Table 1. Cut-off T-scores for the validity scales.*

|  | L | F | K | TRIN | VRIN | Fb | Fp | S |
|---|---|---|---|---|---|---|---|---|
| Profile likely invalid when $T$ | > 64 | >64 | >64 | >80 | >80 | >80 | >100 | >80 |
|  | <40 |  | <40 |  |  |  |  |  |

*Research design*. In this exploratory study, no experimental manipulations were performed. Eventually, participants were to be assigned to clusters (see Statistical Analyses), based on distance between clusters and similarity within clusters.

**Statistical analyses**

To distinguish different personality profiles, I adhered to the protocol outlined by Spaans et al. (2009). I employed hierarchical cluster analysis and *K*-means cluster analysis (see also Spaans et al., 2009). Cluster analysis allows for segmentation of a large set of data into

meaningful groups of similar data, called clusters, based on the clustering variables of interest. In the present study, I used the MMPI-2's Clinical Scales as clustering variables. First, I employed the hierarchical cluster analysis procedure using Ward's method (squared Euclidean distance) to determine the appropriate number of clusters. With this method, all participants are initially treated as separate clusters. Consequently, hierarchical cluster analysis systematically groups data (i.e. participants and, eventually, clusters) that are most identical to each other on the clustering variables of interest. With each step, the within-cluster error sum of squares (ESS) is calculated, and the combination of participants/clusters that provides the least increase in total ESS is chosen. Merging participants/clusters automatically increases ESS. Hierarchical cluster analysis keeps track of the increase in ESS. A relatively high increase in ESS indicates that the clusters that were merged in that particular step might differ in such a way that it is better not to merge these, as the participants within the resulting cluster differ from each other in meaningful ways. The largest increase in ESS usually occurs in the last step of the hierarchical cluster analysis. However, the step before the first sudden jump in ESS indicates the best point to terminate the cluster analysis, regardless of where that first sudden jump occurs (Spaans et al., 2009). I would like to mention here that the decision on the number of clusters is up to the researcher and is therefore necessarily subjective (Spaans et al., 2009; Verma, 2013). The decision to terminate clustering is based on whether the cluster solution provides the researcher with sufficient adequate information in order to proceed. Former studies (e.g. Spaans et al., 2009) compensated for this subjectivity by including multiple solutions, such as 2-, 3-, and 4-cluster solutions. In the present study, different cluster solutions will also be presented.

Finally, distinct groups (clusters) emerge that have the lowest possible within-cluster ESS, while differing significantly from the other clusters. Second, after determining the

appropriate number of clusters, $K$-means cluster analysis based on this fixed number of clusters provided content information on the clusters. It provided the $T$-scores for the ten Clinical Scales per cluster. For each cluster, the L, F, K, TRIN, VRIN, Fb, Fp and S scores were computed to check for profile validity. Independent samples t-tests were performed to compare the mean T-scores of the female and male subpopulation per Clinical Scale in all of the presented clusters.

## Results

Below, the results of the hierarchical cluster analysis and $K$-means cluster analysis are presented first, for the total, female and male population. Next, an alternative cluster solution is discussed. Finally, the results of the independent samples t-tests comparing the female and male population are presented.

### Total population

First, the hierarchical cluster analysis was performed on the total population. The results of this analysis are shown in Figure 1 (last 20 steps of segmentation).

**Figure 1. Hierarchical cluster analysis - total population**
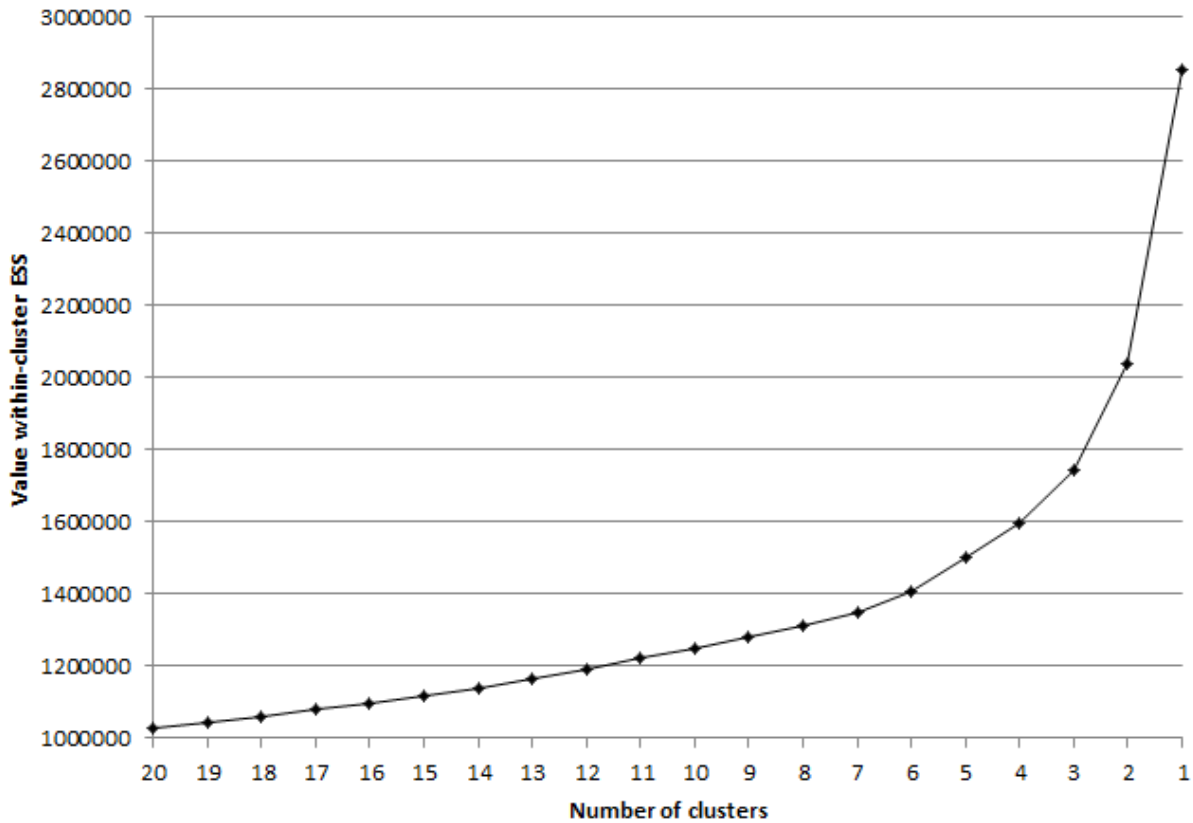


The most marked first increase in within-cluster error sum of squares (ESS) occurred between

cluster 3 and cluster 2. Therefore, a 3-cluster solution was chosen. Subsequently, the *K*-means

cluster analysis, based on the premise of these three clusters, revealed the following three

profiles (see Figure 2; for an overview of the *T*-scores, see Table 2).

Figure 2. MMPI-2 profiles for the 3-cluster solution - total population
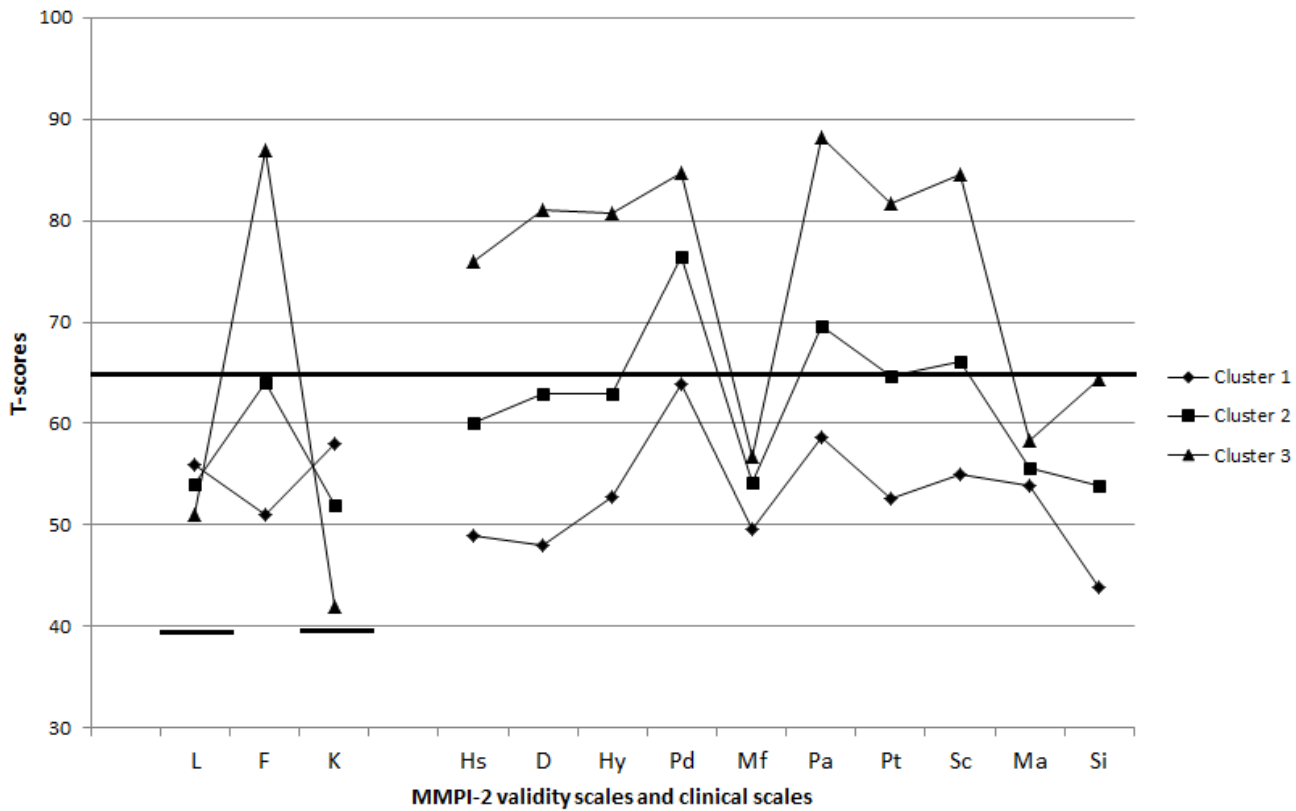
Cluster 1 (49%, $n = 817$) consisted of participants who reported $T$-scores ranging from 44 to 64

(i.e., no clinical elevations on any of the clinical scales). The shape of the L-F-K scale pattern

(downward arrow) indicates a tendency to underreport personal/emotional problems (see ).

Viewed separately, the $T$-scores for the L, F, and K scales all remained within the normal range,

indicating a valid clinical scale profile for cluster 1 (see Table 3). Cluster 2 (36%, $n = 602$)

showed a distinct clinical profile with elevations on the Psychopathic Deviate, Paranoia,

*Table 2. The 3-cluster solution in the total, female, and male populations: Clusters, mean T-scores, and results of the independent samples t-tests.*

| | N (%) | Hs | D | Hy | Pd | Mf | Pa | Pt | Sc | Ma | Si |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Cluster 1** | | | | | | | | | | | |
| Total | 817 (49) | 49 | 48 | 53 | 64 | 50 | 59 | 53 | 55 | 54 | 44 |
| Female | 109 (44) | 49 | 48 | 51 | 61 | 53 | 57 | 51 | 54 | 53 | 44 |
| Male | 708 (50) | 49 | 48 | 53 | 64 | 49 | 59 | 53 | 55 | 54 | 44 |
| $t$ (female-male)[a] | | -1.35 | -.20 | -3.30 | -2.01 | 2.78 | -1.10 | -3.18 | -.66 | .34 | .24 |
| df[a] | | 815 | 815 | 815 | 815 | 133 | 815 | 815 | 815 | 815 | 815 |
| $p$-value[a] | | .176 | .845 | **.001** | **.045** | **.006** | .272 | **.002** | .513 | .736 | .807 |
| **Cluster 2** | | | | | | | | | | | |
| Total | 602 (36) | 60 | 63 | 63 | **76** | 54 | **70** | **65** | **66** | 56 | 54 |
| Female | 80 (32) | 58 | 62 | 60 | **77** | 52 | **74** | **65** | **71** | 59 | 54 |
| Male | 522 (37) | 59 | 62 | 63 | **76** | 54 | **69** | 64 | **65** | 55 | 53 |
| $t$ (female-male)[a] | | -.52 | 1.03 | -1.22 | -.86 | -3.06 | 2.89 | 1.34 | 4.51 | .24 | 1.95 |
| df[a] | | 600 | 600 | 600 | 600 | 600 | 600 | 97 | 94 | 600 | 600 |
| $p$-value[a] | | .603 | .302 | .222 | .393 | **.002** | **.004** | .184 | **<.001** | .813 | .052 |
| **Cluster 3** | | | | | | | | | | | |
| Total | 244 (15) | **76** | **81** | **81** | **85** | 57 | **88** | **82** | **84** | 58 | 64 |
| Female | 60 (24) | **74** | **82** | **78** | **79** | 49 | **88** | **82** | **89** | 56 | **65** |
| Male | 184 (13) | **76** | **81** | **81** | **86** | 59 | **87** | **81** | **82** | 59 | 64 |
| $t$ (female-male)[a] | | -1.06 | -.11 | -1.57 | -3.86 | -5.21 | 1.60 | .39 | 4.52 | .18 | .37 |
| df[a] | | 242 | 242 | 242 | 242 | 242 | 242 | 85 | 79 | 242 | 242 |
| $p$-value[a] | | .290 | .915 | .118 | **<.001** | **<.001** | .111 | .697 | **<.001** | .858 | .709 |

*Note.* Clinically elevated *T*-scores and significant *p*-values ($\alpha$, two-tailed = .05) are printed in boldface. Hs = Hypochondriasis; D = Depression; Hy = Hysteria; Pd = Psychopathic Deviate; Mf = Masculinity/femininity; Pa = Paranoia; Pt = Psychasthenia, Sc = Schizophrenia, Ma = Hypomania; Si = Social introversion; *t* = t-test score; df = degrees of freedom.

[a]The independent samples t-tests (*t*, df, and *p*-value) tested the null hypothesis of equal mean *T*-scores for the female versus male population, per cluster.

Psychasthenia, and Schizophrenia scales (a 4-6-7-8 profile). The L-F-K scale pattern (upward arrow) reveals that the cluster 2 population generally admitted to having personal/emotional problems. Furthermore, the T-scores for the L, F, and K scales did not exceed their respective cut-off points, meaning the clinical scale profile was valid. Cluster 3 (15%, $n = 244$) showed clinical elevations on the Hypochondriasis, Depression, Hysteria, Psychopathic Deviate, Paranoia, Psychasthenia, and Schizophrenia scales (a 1-2-3-4-6-7-8 profile). On six out of eight elevated scales, $T$-scores above 80 were reported. The L-F-K scale pattern (upward arrow) pointed to a general tendency to admit personal/emotional problems. The $T$-score of 87 on the F-scale seemed to indicate an invalid clinical scale profile. The TRIN ($T = 58$), VRIN ($T = 55$), Fp ($T = 67$), and S ($T = 37$) scores fell within their respective normal ranges. However, Fb ($T = 92$) was elevated. The F and Fb scales indicate an exaggeration of personal problems in cluster 3.

*Table 3. T-scores for the validity scales in the 3-cluster solution (total, female, and male populations). Clinically elevated T-scores are printed in boldface.*

|  | L | F | K | TRIN | VRIN | Fb | Fp | S |
|---|---|---|---|---|---|---|---|---|
| Cluster 1 |  |  |  |  |  |  |  |  |
| Total | 56 | 51 | 58 | 56 | 46 | 51 | 53 | 54 |
| Female | 56 | 51 | 57 | 56 | 46 | 50 | 56 | 55 |
| Male | 56 | 51 | 58 | 56 | 46 | 51 | 52 | 54 |
| Cluster 2 |  |  |  |  |  |  |  |  |
| Total | 54 | 64 | 52 | 59 | 53 | 65 | 57 | 47 |
| Female | 54 | **71** | 48 | 58 | 53 | 68 | 63 | 43 |
| Male | 54 | 62 | 53 | 59 | 53 | 64 | 56 | 48 |
| Cluster 3 |  |  |  |  |  |  |  |  |
| Total | 51 | **87** | 42 | 58 | 55 | **92** | 67 | 37 |
| Female | 52 | **87** | 42 | 56 | 56 | **82** | 70 | 38 |
| Male | 51 | **86** | 43 | 59 | 55 | **91** | 65 | 37 |

*Note*. Clinically elevated *T*-scores are printed in boldface. L = Lie; F = Infrequency; K = Correction; TRIN = True Response Inconsistency; VRIN = Variable Response Inconsistency; Fb = F-back; Fp = Infrequency-Psychopathology; S = Superlative self-presentation.

**Female population**

Again, the hierarchical cluster analysis was performed first. Figure 3 shows the last 20 steps of segmentation by the hierarchical cluster analysis for the female population.



Figure 3. Hierarchical cluster analysis - female population
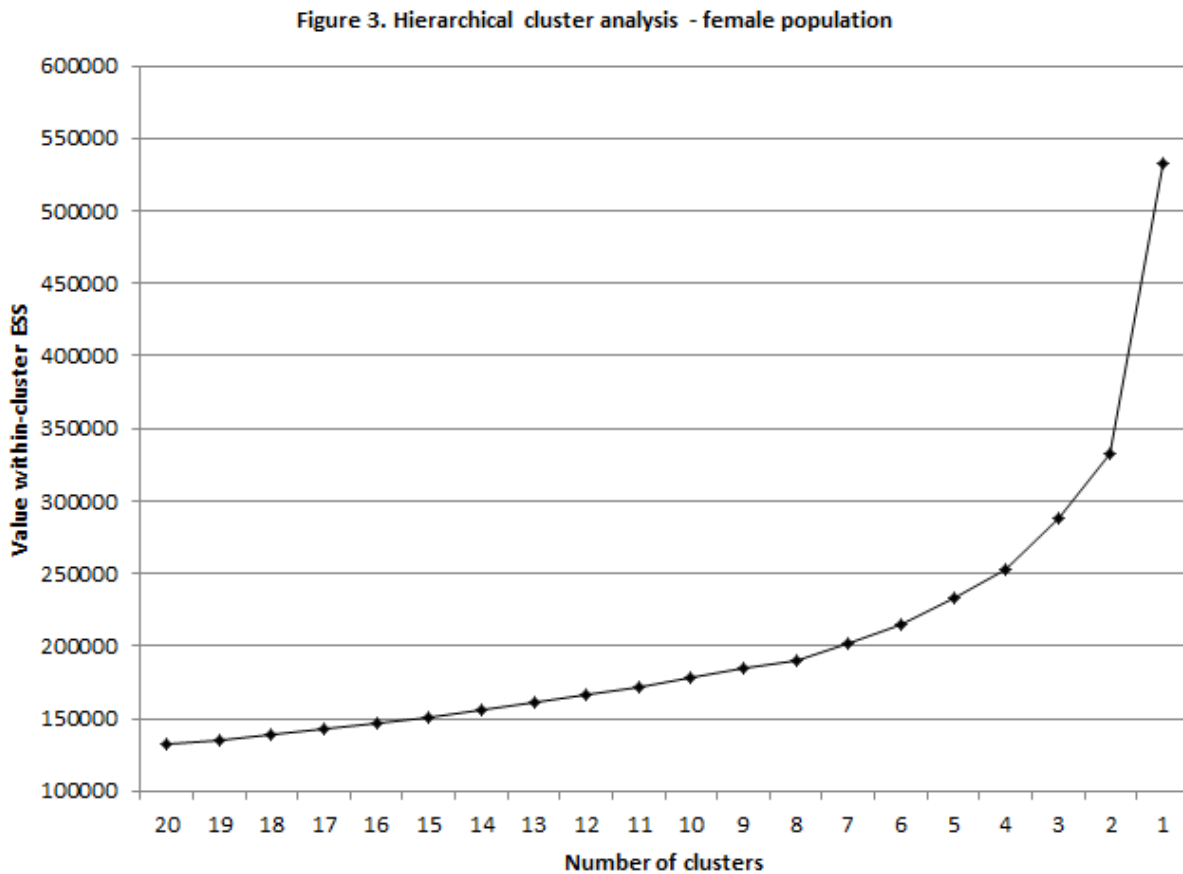
The ESS curve increased gradually, and offered a less straightforward point of demarcation compared to the curve for the total population. Between cluster 2 and cluster 1, the ESS 'jumped' up most markedly for the first time, indicating that a 2-cluster solution was the most preferable outcome (see Figure 4; for $T$-scores, see Table 4). Cluster 1 (39%, $n = 98$) showed a '1-2-3-4-6-7-8' profile: clinically elevated $T$-scores emerged on Hypochondriasis, Depression, Hysteria, Psychopathic Deviate, Paranoia, Psychasthenia, and Schizophrenia. According to the pattern of

the L-F-K scales, the female subpopulation in cluster 1 acknowledged personal/emotional problems. The high $T$-score (82) on the F-scale warranted for further scrutiny of the validity scales. L ($T = 52$), K ($T = 43$), TRIN ($T = 57$), VRIN ($T = 55$), Fp ($T = 67$), and S ($T = 39$) scores all fell within average range (see Table 5). However, Fb ($T = 86$) was elevated. The F and Fb scale revealed an exaggeration of personal and emotional problems in cluster 1.



Figure 4. MMPI-2 profiles for the 2-cluster solution - female population
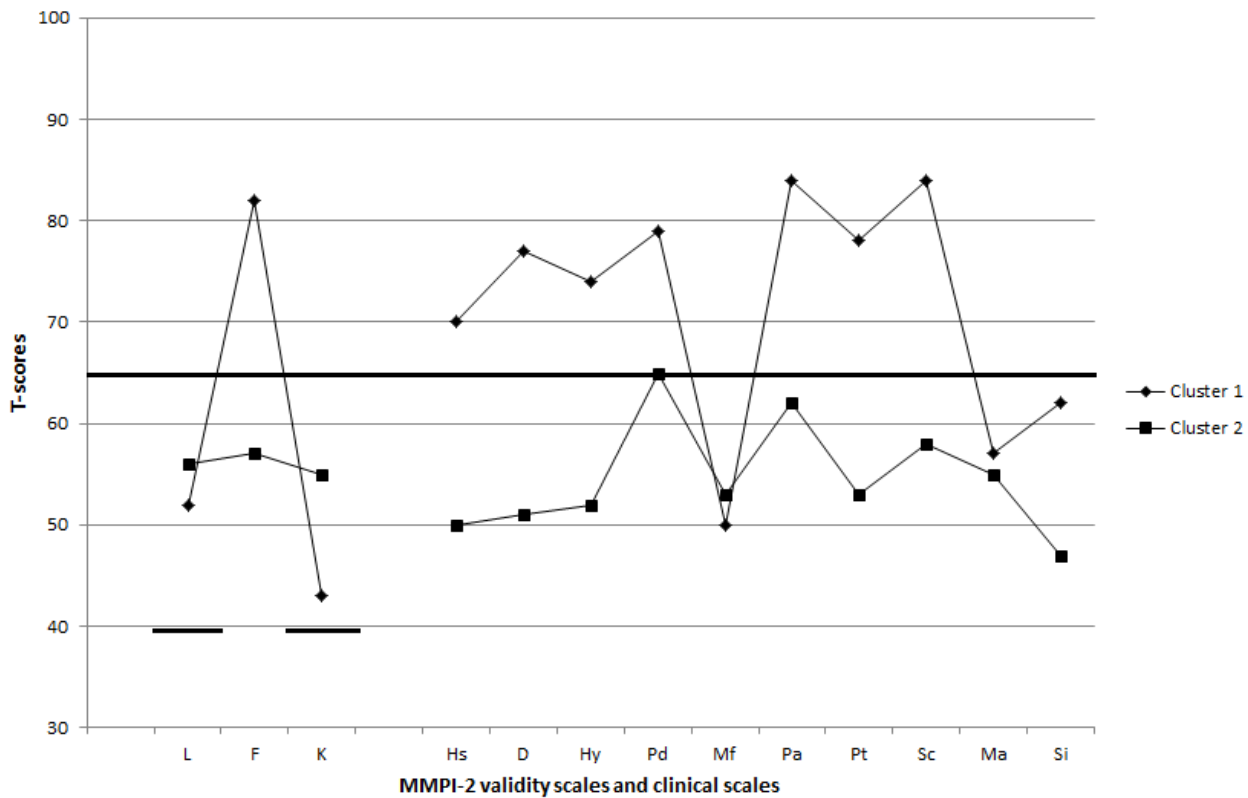
In Cluster 2 (61%, $n = 151$), only on the Psychopathic Deviate (4) scale the cut-off $T$-score of 65 was obtained; the lowest $T$-score (47) for this cluster was found on Social Introversion. The clinical profile was valid (i.e., all validity scales showed $T$-scores within average range; see Table 5).

20

Table 4. The 2-cluster solution for the female and male populations: Clusters, T-scores, and results of the independent samples t-tests.

| | N (%) | Hs | D | Hy | Pd | Mf | Pa | Pt | Sc | Ma | Si |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Cluster 1** | | | | | | | | | | | |
| Female | 98 (39) | **70** | **77** | **74** | **79** | 50 | **84** | **78** | **84** | 57 | 62 |
| Male | 379 (27) | **70** | **75** | **74** | **83** | 57 | **80** | **75** | **76** | 58 | 60 |
| $t$ (female-male)[a] | | .33 | 1.58 | -.03 | -3.40 | -6.49 | 2.56 | 2.31 | 4.89 | -.75 | 1.42 |
| df[a] | | 475 | 475 | 180 | 475 | 475 | 475 | 475 | 122 | 475 | 475 |
| $p$-value[a] | | .745 | .115 | .974 | **.001** | **<.001** | **.011** | **.021** | **<.001** | .456 | .158 |
| **Cluster 2** | | | | | | | | | | | |
| Female | 151 (61) | 50 | 51 | 52 | **65** | 53 | 62 | 53 | 58 | 55 | 47 |
| Male | 1035 (73) | 52 | 52 | 56 | **67** | 51 | 61 | 56 | 57 | 54 | 46 |
| $t$ (female-male)[a] | | -2.12 | -1.01 | -4.40 | -2.06 | 2.21 | .21 | -3.14 | 1.08 | 1.27 | .67 |
| df[a] | | 1184 | 1184 | 1184 | 1184 | 1184 | 180 | 1184 | 173 | 1184 | 1184 |
| $p$-value[a] | | **.034** | .314 | **<.001** | **.040** | **.027** | .835 | **.002** | .281 | .204 | .505 |

*Note*. Clinically elevated *T*-scores and significant *p*-values (α, two-tailed = .05) are printed in boldface. Hs = Hypochondriasis; D = Depression; Hy = Hysteria; Pd = Psychopathic Deviate; Mf = Masculinity/femininity; Pa = Paranoia; Pt = Psychasthenia, Sc = Schizophrenia, Ma = Hypomania; Si = Social introversion; *t* = *t*-test score; df = degrees of freedom.

[a]The independent samples t-tests (*t*, df, and *p*-value) tested the null hypothesis of equal mean *T*-scores for the female versus male population, per cluster.

Table 5. T-scores for the validity scales in the 2-cluster solution (total, female, and male populations).

| | L | F | K | TRIN | VRIN | Fb | Fp | S |
|---|---|---|---|---|---|---|---|---|
| **Cluster 1** | | | | | | | | |
| Female | 52 | **82** | 43 | 57 | 55 | **86** | 67 | 39 |
| Male | 52 | **77** | 46 | 59 | 56 | **81** | 62 | 41 |
| **Cluster 2** | | | | | | | | |
| Female | 56 | 57 | 55 | 57 | 48 | 54 | 58 | 52 |
| Male | 55 | 54 | 57 | 57 | 48 | 54 | 53 | 52 |

*Note*. Clinically elevated *T*-scores are printed in boldface. L = Lie; F = Infrequency; K = Correction; TRIN = True Response Inconsistency; VRIN = Variable Response Inconsistency; Fb = F-back; Fp = Infrequency-Psychopathology; S = Superlative self-presentation.
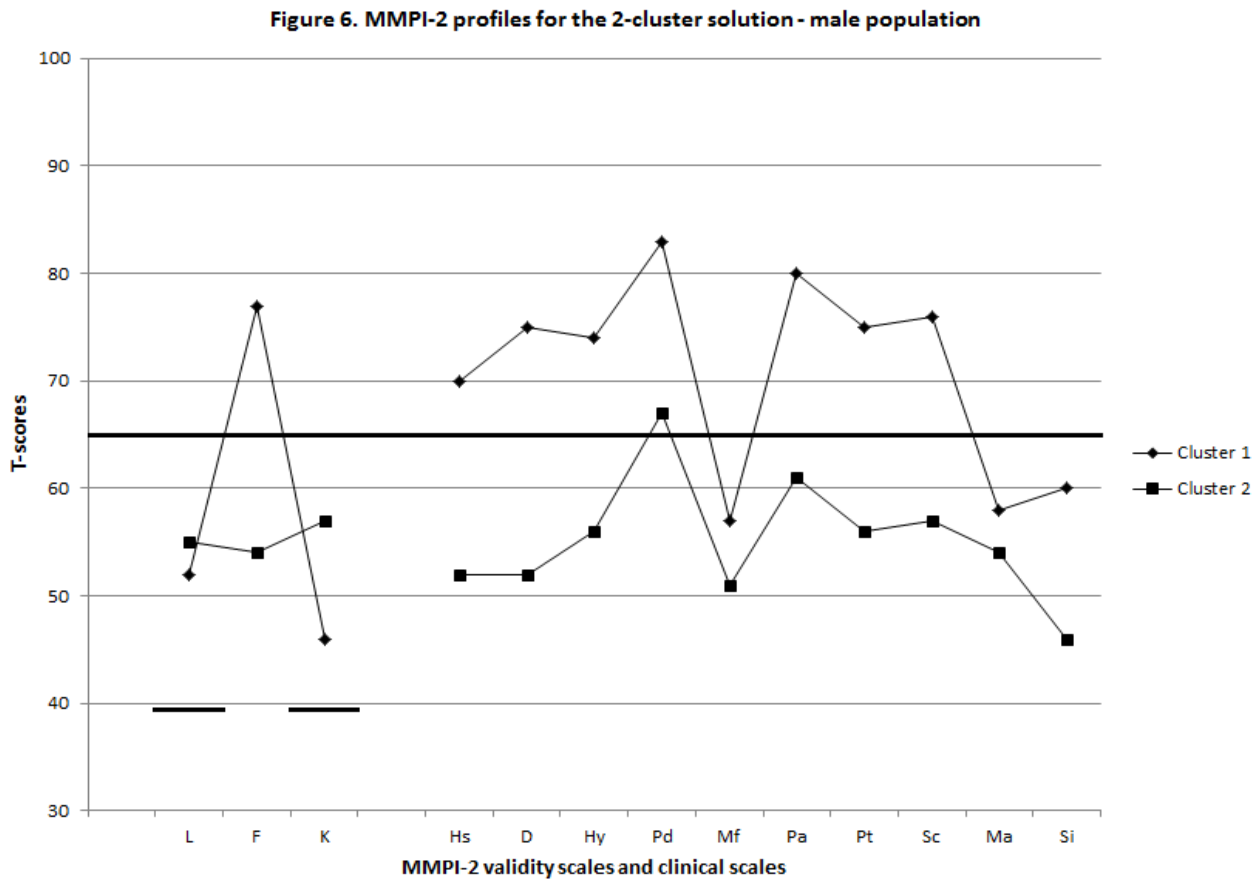
**Male population**

Figure 5 shows the results of the hierarchical cluster analysis for the male population.



Figure 5. Hierarchical cluster analysis - male population
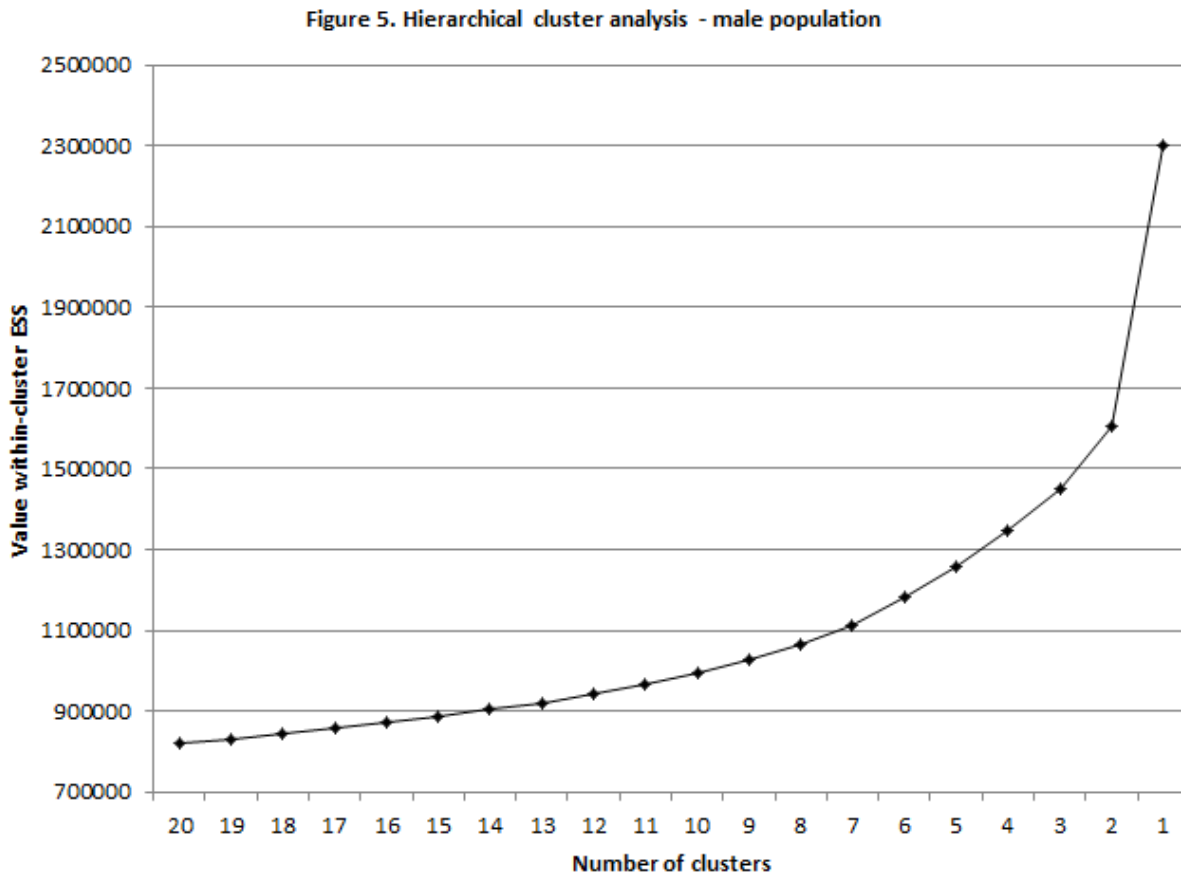
Again, a 2-cluster solution was chosen as the ESS curve did not show a marked increase before the segmentation between the last two clusters. Figure 6 shows the results of the $K$-means analysis (see also Table 4). The first cluster (27%, $n = 379$) was formed by clinically elevated $T$-scores on the Hypochondriasis, Depression, Hysteria, Psychopathic Deviate, Paranoia, Psychasthenia, and Schizophrenia scales: the '1-2-3-4-6-7-8' profile. The shape of the L-F-K-pattern revealed participant's acknowledgement of personal problems. The $T$-score of 77 on the $F$-scale could represent a modest exaggeration of reported symptoms (see Table 5). The scores

on L ($T = 52$), K ($T = 46$), TRIN ($T = 59$), VRIN ($T = 56$), Fp ($T = 62$), and S ($T = 41$) fell in the

normal range. The Fb ($T = 81$) scale, together with the F-scale, indicated overreporting of

perceived symptoms by cluster 1. The second cluster (73%, $n = 1035$) of the male subpopulation

contained a valid '4' profile: a clinically elevated score on Psychopathic Deviate ($T = 67$) was

found.



Figure 6. MMPI-2 profiles for the 2-cluster solution - male population

**4-cluster solution**

A 4-cluster solution was included for the total populations, and separately for the female and

male subpopulations. Hierarchical cluster analysis was not performed as the number of clusters

was fixed. The results of the $K$-means cluster analysis are shown in Table 6, $T$-scores for the

validity scales are shown in Table 7.

23

*Table 6. The 4-cluster solution in the total, female, and male populations: Clusters, T-scores, and results of the independent samples t-tests.*

| | N (%) | Hs | D | Hy | Pd | Mf | Pa | Pt | Sc | Ma | Si |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Cluster 1** | | | | | | | | | | | |
| Total | 279 (17) | 53 | 53 | 54 | **77** | 52 | **71** | 60 | **67** | **72** | 47 |
| Female | 28 (11) | 55 | 54 | 54 | **78** | 58 | **79** | 64 | **80** | **77** | 49 |
| Male | 472 (33) | 59 | 61 | 63 | **73** | 54 | 64 | 62 | 62 | 49 | 52 |
| *t* (female-male)[a] | | -2.11 | -4.07 | -4.80 | 2.18 | 1.39 | 5.98 | 1.53 | 8.33 | 11.95 | -1.28 |
| df[a] | | 498 | 498 | 498 | 498 | 28 | 28 | 498 | 28 | 29 | 498 |
| *p*-value[a] | | **.035** | **<.001** | **<.001** | **.029** | .175 | **<.001** | .127 | **<.001** | **<.001** | .203 |
| **Cluster 2** | | | | | | | | | | | |
| Total | 464 (28) | 62 | **66** | **66** | 75 | 55 | **69** | 65 | 65 | 50 | 55 |
| Female | 68 (27) | 64 | **70** | **68** | 75 | 47 | **74** | 68 | 69 | 50 | 58 |
| Male | 203 (14) | 60 | **65** | 61 | **83** | 56 | **80** | 70 | 74 | 68 | 56 |
| *t* (female-male)[a] | | 2.79 | 4.16 | 4.10 | -5.20 | -6.16 | -3.58 | -1.59 | -4.50 | -11.90 | 1.33 |
| df[a] | | 269 | 269 | 269 | 269 | 269 | 269 | 269 | 269 | 203 | 269 |
| *p*-value[a] | | **.006** | **<.001** | **<.001** | **<.001** | **<.001** | **<.001** | .112 | **<.001** | **<.001** | .183 |
| **Cluster 3** | | | | | | | | | | | |
| Total | 215 (13) | **76** | **82** | **81** | **86** | 57 | **89** | **83** | **86** | 59 | **66** |
| Female | 43 (17) | **76** | **85** | **79** | **82** | 51 | **92** | **86** | **96** | 58 | **69** |
| Male | 142 (10) | **80** | **85** | **86** | **86** | 59 | **86** | **82** | **82** | 55 | **66** |
| *t* (female-male)[a] | | -2.05 | .06 | -3.38 | -2.27 | -4.54 | 2.63 | 2.19 | 7.68 | 1.13 | 1.68 |
| df[a] | | 183 | 183 | 183 | 183 | 183 | 183 | 183 | 183 | 183 | 183 |
| *p*-value[a] | | **.042** | .951 | **.001** | **.024** | **<.001** | **.009** | **.030** | **<.001** | .261 | .094 |
| **Cluster 4** | | | | | | | | | | | |
| Total | 705 (42) | 49 | 48 | 53 | 62 | 50 | 58 | 52 | 54 | 51 | 44 |
| Female | 110 (45) | 49 | 48 | 51 | 62 | 53 | 58 | 51 | 55 | 53 | 45 |
| Male | 597 (43) | 48 | 47 | 52 | 63 | 49 | 59 | 52 | 54 | 55 | 43 |
| *t* (female-male)[a] | | .37 | 1.94 | -1.58 | -1.44 | 4.32 | -.90 | -1.80 | .22 | -1.84 | 1.47 |
| df[a] | | 705 | 705 | 705 | 705 | 705 | 705 | 705 | 705 | 173 | 136 |
| *p*-value[a] | | .714 | .052 | .114 | .150 | **<.001** | .366 | .072 | .830 | .067 | .143 |

*Note*. Clinically elevated *T*-scores and significant *p*-values ($\alpha$, two-tailed = .05) are printed in boldface. Hs = Hypochondriasis; D = Depression; Hy = Hysteria; Pd = Psychopathic Deviate; Mf = Masculinity/femininity; Pa = Paranoia; Pt = Psychasthenia, Sc = Schizophrenia, Ma = Hypomania; Si = Social introversion; *t* = t-test score; df = degrees of freedom.

[a]The independent samples t-tests (*t*, df, and *p*-value) tested the null hypothesis of equal mean *T*-scores for the female versus male population, per cluster.

*Table 7. T-scores for the validity scales in the 4-cluster solution (total, female, and male populations).*

|  | L | F | K | TRIN | VRIN | Fb | Fp | S |
|---|---|---|---|---|---|---|---|---|
| **Cluster 1** | | | | | | | | |
| Total | 51 | **68** | 47 | 60 | 54 | 67 | 61 | 42 |
| Female | 53 | **84** | 42 | 59 | 56 | **81** | 77 | 35 |
| Male | 56 | 56 | 57 | 58 | 50 | 58 | 53 | 52 |
| **Cluster 2** | | | | | | | | |
| Total | 55 | 61 | 53 | 58 | 52 | 63 | 55 | 49 |
| Female | 56 | **66** | 49 | 57 | 54 | 66 | 59 | 47 |
| Male | 49 | **78** | 44 | 59 | 57 | 78 | 63 | 38 |
| **Cluster 3** | | | | | | | | |
| Total | 51 | **90** | 42 | 58 | 55 | **95** | 68 | 36 |
| Female | 51 | **94** | 40 | 57 | 54 | **100** | 73 | 35 |
| Male | 52 | **84** | 44 | 59 | 54 | **92** | 63 | 39 |
| **Cluster 4** | | | | | | | | |
| Total | 56 | 50 | 59 | 55 | 46 | 50 | 52 | 56 |
| Female | 55 | 52 | 57 | 56 | 46 | 51 | 56 | 55 |
| Male | 56 | 51 | 57 | 56 | 46 | 51 | 52 | 53 |

*Note*. Clinically elevated *T*-scores are printed in boldface. L = Lie; F = Infrequency; K = Correction; TRIN = True Response Inconsistency; VRIN = Variable Response Inconsistency; Fb = F-back; Fp = Infrequency-Psychopathology; S = Superlative self-presentation.

Cluster 1, in the total population (17%, $n = 279$), revealed a '4-6-7-8' profile with elevated scores on Psychopathic deviate, Paranoia, Schizophrenia and Hypomania. A modest exaggeration of experienced problems was indicated by the F scale ($T = 68$). As all other validity scales fell within acceptable range, the profile was still considered valid. The female subpopulation in cluster 1 (11%, $n = 28$) revealed the same profile as the total population; however, this population appeared to over report personal/emotional problems ($T$-scores of 84 and 81 for F and Fb, respectively). The male subpopulation (33%, $n = 472$) only had an elevated score on Psychopathic deviate ($T = 73$); the profile was valid.

Cluster 2 revealed a '2-3-4-6-7-8' profile for the total population (28%, $n = 464$), which was reflected in the female subpopulation (27%, $n = 68$). Elevated scores were reported on

Depression, Hysteria, Psychopathic deviate, Paranoia, Psychasthenia, and Schizophrenia.The

profile for the total and female population was valid; the other validity scales compensated for the

modestly elevated T-score on the F scale ($T = 66$). The male subpopulation (14%, $n = 203$) in

cluster 2 showed a somewhat different clinical profile, with elevated scores on Depression,

Psychopathic deviate, Paranoia, Psychasthenia, Schizophrenia, and Hypomania (a '2-4-6-7-8-9'

profile). An elevated score on the F scale ($T = 78$) indicated a modest exaggeration of symptoms

for this population.

Cluster 3 consisted of participants who reported the most psychological distress. A '1-2-3-

4-6-7-8-0' profile emerged for the total population (13%, $n = 215$), and for the female (17%, $n =$

43) and male (10%, $n = 142$) subpopulations. Elevated scores were reported on Hypochondriasis,

Depression, Hysteria, Psychopathic deviate, Paranoia, Psychasthenia, Schizophrenia, and Social

introversion. In the female subpopulation, on three Clinical Scales $T$-scores above 80 were

reported and two Clinical Scales had $T$-scores above 90. In the male subpopulation, on seven out

of eight Clinical Scales, participants obtained $T$-scores of 80 or more. However, further analysis

of the validity scales revealed that participants in cluster 3 could have been heavily exaggerating

their symptoms. Strongly elevated scores were found on the F scale for the total population ($T =$

90 ), and for the female ($T = 94$) and male ($T = 84$) subpopulations. Combined with strongly

elevated scores on Fb ($T = 95$, $T = 100$, and $T = 92$ for the total, female, and male population,

respectively), this could indicate a cry for help by the participants in cluster 3.

Cluster 4 showed no clinical elevations on the Clinical Scales for the total population

(42%, $n = 705$), and for the female (45%, $n = 110$) and male (43%, $n = 597$) subpopulations. The

overall relationship between the L, F, and K scales could indicate under reporting of symptoms;

however, $T$-scores for all validity scales feel within normal range. The clinical profile was valid.

**Female - male comparisons**

Comparisons between the female and male subpopulation were performed by employing independent samples t-tests. For both populations in all of the presented cluster solutions, the t-tests were based on the mean $T$-scores per Clinical Scale and per cluster.

*2-cluster solution.* First, the female and male populations were compared on the 2-cluster solution (see Table 4) with independent samples t-tests. In both populations, cluster 1 comprised a '1-2-3-4-6-7-8' profile. Within this cluster, the male population had a significantly higher $T$-score on scale 4 - Psychopathic Deviate ($T = 83$, $t(475) = -3.40$, $p = .001$), as well as on scale 5 - Masculinity/Femininity ($T = 57$, $t(475) = -6.49$, $p < .001$). The female population scored significantly higher on scale 6 - Paranoia ($T = 84$, $t(475) = 2.56$, $p = .011$), scale 7 - Psychasthenia ($T = 78$, $t(475) = 2.31$, $p = .021$) and scale 8 - Schizophrenia ($T = 84$, $t(122) = 4.89$, $p < .001$). For cluster 2, only scale 4 showed an elevated $T$-score for both men ($T = 67$) and women ($T = 65$) and this difference was statistically significant ($t(1184) = -3.40$, $p = .040$). However, significant differences were also found on four scales that had no elevated $T$-scores. Women scored significantly higher than men on Masculinity/Femininity ($T = 53$, $t(1184) = 2.21$, $p = .027$); men scored significantly higher on Hypochondriasis ($T = 52$, $t(1184) = -2.12$, $p = .034$), Hysteria ($T = 56$, $t(1184) = -4.40$, $p < .001$) and Psychasthenia ($T = 56$, $t(1184) = -3.14$, $p = .002$).

*3-cluster solution.* To investigate an alternative 3-cluster solution for, and to allow for further comparisons between, the female and male populations, the three clusters of the total population clusters were split up. The results of the $K$-means cluster analysis are included in Table 2. In the clinically 'healthy' first cluster, the male population ($50\%$, $n = 708$) showed a profile of higher $T$-scores than female population ($44\%$, $n = 109$) on scale 3 ($T = 53$, $t(815) = -3.30$, $p = .001$), scale 4 ($T = 64$, $t(815) = -2.01$, $p = .045$), and scale 7 ($T = 53$, $t(815) = -3.18$, $p = .002$), whereas women scored higher on scale 5 ($T = 53$, $t(133) = 2.78$, $p = .006$). In cluster 2, the

total population '4-6-7-8' profile was reflected in cluster 2 of the female population (32% $n =$ 80); the male population (37%, $n = 522$) in this cluster was marked by a '4-6-8' profile. The female population scored significantly higher than the male population on scales 6 ($T = 74$, $t(600) = 2.89$, $p = .004$) and 8 ($T = 71$, $t(94) = 4.51$, $p < .001$), whereas the male population scored higher on scale 5 ($T = 54$, $t(600) = -3.06$, $p = .002$). In cluster 3, the male population (13%, $n = 184$) reported a '1-2-3-4-6-7-8' profile with six out of eight $T$-scores above 80. The female population in cluster 3 (24%, $n = 60$) reported a '1-2-3-4-6-7-8-0' profile, with four $T$-scores above 80. The female population scored significantly higher than the male population on scale 8 ($T = 89$, $t(79) = 4.52$, $p < .001$). The male population scored significantly higher than the female population on scale 4 ($T = 86$, $t(242) = -3.86$, $p < .001$) and scale 5 ($T = 59$, $t(242) = -5.21$, $p < .001$).

*4-cluster solution.* The various profiles have been presented earlier (see 4-cluster solution above). The results of the independent samples t-tests were included in Table 6. In cluster 1, the female population scored higher on scale 4 ($T = 78$, $t(498) = 2.18$, $p = .029$), scale 6 ($T = 79$, $t(28) = 5.98$, $p < .001$), scale 8 ($T = 80$, $t(28) = 8.33$, $p < .001$), and scale 9 ($T = 77$, $t(29) = 11.95$, $p < .001$) than the male population. The male population in this cluster scored higher on scale 1 ($T = 59$, $t(498) = -2.11$, $p = .035$), scale 2 ($T = 61$, $t(498) = -4.07$, $p < .001$), and scale 3 ($T = 63$, $t(498) = -4.80$, $p < .001$). In cluster 2, the two subpopulations differed significantly on scales, 1, 2, 3, 4, 5, 6, 8, and 9. The female population scored significantly higher on scale 1 ($T = 64$, $t(269) = 2.79$, $p = .006$), scale 2 ($T = 70$, $t(269) = 4.16$, $p < .001$), and scale 3 ($T = 68$, $t(269) = 4.10$, $p < .001$) than the male population. The male population showed significantly higher scores on scale 4 ($T = 83$, $t(269) = -5.20$, $p < .001$), scale 5 ($T = 56$, $t(269) = -6.16$, $p < .001$), scale 6 ($T = 80$, $t(269) = -3.58$, $p < .001$), scale 8 ($T = 74$, $t(269) = -4.50$, $p < .001$), and scale 9 ($T = 68$, $t(203) = -11.90$, $p < .001$) than the female population. Cluster 3 represented participants who reported the most

significant psychological distress, compared to the other three clusters. As mentioned earlier, this might be a highly exaggerated symptom profile. The female population scored significantly higher on scale 6 ($T = 92$, $t(183) = 2.63$, $p = .009$), scale 7 ($T = 86$, $t(183) = 2.19$, $p = .030$), and scale 8 ($T = 96$, $t(183) = 7.68$, $p < .001$) than the male population. The male population showed significantly higher scores on scale 1 ($T = 80$, $t(183) = -2.05$, $p < .001$), scale 3 ($T = 86$, $t(183) = -3.38$, $p < .001$), scale 4 ($T = 86$, $t(183) = -2.27$, $p < .001$), and scale 5 ($T = 86$, $t(183) = -4.54$, $p < .001$) than the female population. Cluster 4 only showed a significantly higher score for the female subpopulation on scale 5 ($T = 53$, $t(705) = 4.32$, $p < .001$).

Table 8 provides a summary of the profile codes for the total population and for the female and male subpopulation, as well as the comparisons for the female versus male subpopulation, for each cluster solution.

## Discussion

In the present study, the focus was on whether self-reported use of the MMPI-2 can successfully distinguish various personality profiles in forensic pre-trial settings. For the general forensic population, a 3-cluster solution seemed to be the most preferable. This solution resulted in a non-disturbed (no clinical elevations), mildly disturbed ('4-6-7-8') , and very disturbed ('1-2-3-4-6-7-8' profile) cluster. The first research hypothesis was thereby not confirmed, although it must be mentioned that the present study did not yield such elaborate distinct MMPI-2 profiles found by previous studies (e.g. Nieberding et al., 2003; and Megargee et al., 1979, 2001). The MMPI-2 was successful in discerning various personality profiles in a general population, thereby providing a positive answer to the main research question. Focused on gender, the mildly disturbed cluster of the total population was reflected in the female population; for the male

*Table 8. Profile codes for the 2-, 3-, and 4-cluster solution per (sub)population.*

| | Cluster 1 | Cluster 2 | Cluster 3 | Cluster 4 |
|---|---|---|---|---|
| **2-cluster solution** | | | | |
| Female[a] | 1-2-3-4-6-7-8[c] | 4 | | |
| Male[a] | 1-2-3-4-6-7-8[c] | 4 | | |
| Female > Male[b] | 6-7-8 | 5 | | |
| Male > Female[b] | 4-5 | 1-3-4-7 | | |
| **3-cluster solution** | | | | |
| Total[a] | - | 4-6-7-8 | 1-2-3-4-6-7-8[c] | |
| Female[a] | - | 4-6-7-8 | 1-2-3-4-6-7-8-0[c] | |
| Male[a] | - | 4-6-8 | 1-2-3-4-6-7-8[c] | |
| Female > Male[b] | 3-4-7 | 6-8 | 8 | |
| Male > Female[b] | 5 | 5 | 4-5 | |
| **4-cluster solution** | | | | |
| Total[a] | 4-6-8-9 | 2-3-4-6-7-8 | 1-2-3-4-6-7-8-0[c] | - |
| Female[a] | 4-6-8-9[c] | 2-3-4-6-7-8 | 1-2-3-4-6-7-8-0[c] | - |
| Male[a] | 4 | 2-4-6-7-8-9 | 1-2-3-4-6-7-8-0[c] | - |
| Female > Male[b] | 4-6-8-9 | 1-2-3 | 6-7-8 | 5 |
| Male > Female[b] | 1-2-3 | 4-5-6-8-9 | 1-3-4-5 | - |

*Note*. Profile code numbers represent the Clinical Scales: 1 = Hypochondriasis; 2 = Depression; 3 = Hysteria; 4 = Psychopathic Deviate;

5 = Masculinity/femininity; 6 = Paranoia; 7 = Psychasthenia; 8 = Schizophrenia; 9 = Hypomania; 0 = Social introversion.

[a] Profile codes incorporating the elevated mean *T*-scores.

[b] Codes of the scales more strongly reported by men/women than by women/men.

[c] Clinical profiles possibly invalid due to over reporting/exaggeration of symptoms.

population, scale 7 was not clinically elevated. The very disturbed cluster in the female population also included scale '0'. Comparing this study's 3-cluster solution with Spaans et al.'s (2009) 3-cluster solution, identical clusters for the male subpopulation were found.

For both the female and male subpopulations, a 2-cluster solution emerged as the most preferable solution. The two clusters, for both genders, entailed a very disturbed and a non-disturbed cluster. Spaans et al. (2009), focusing solely on a male subpopulation, also found a cluster reporting a disturbed profile and a non-disturbed profile. Scale 4 'Psychopathic deviate', elevated in the 'non-disturbed' labeled cluster in this study and in Spaans et al.'s (2009) study, includes items concerning past contact with the justice system and several other static factors over which the person might not have control anymore. It is, therefore, to be expected that defendant populations generally obtain the minimum cut-off score for scale 4. The findings of Spaans et al. (2009) concerning the 2-cluster solution have been replicated in this study, as both profiles had clinical elevations on the same Clinical Scales. It could therefore also be argued that the first hypothesis has been confirmed, as Spaans et al. (2009) focused solely on a male population. Furthermore, the present study extended these findings to a female subpopulation.

An alternative 4-cluster solution was included in this study as well, to provide a more complete view on the defendant population as the decision on the number of clusters based on hierarchical cluster analysis lacks clear-cut criteria (Spaans et al. 2009; Verma, 2013). This cluster solution yielded a mildly disturbed cluster with a '4-6-8-9' profile (except for the male subpopulation, where a '4' profile was found); a disturbed cluster entailing a '2(-3)-4-6-7-8(-9)' profile; a very disturbed cluster yielding a '1-2-3-4-6-7-8-0' profile; and a non-disturbed cluster with no clinical elevations on any of the Clinical Scales. These findings replicate to a great extent the results of the 4-cluster solution found by Spaans et al. (2009): in a male population, they

found a '4' cluster, a '2-3-4-6-7-8' cluster, a '1-2-3-4-6-7-8' cluster, and a non-disturbed cluster with no clinical elevations on any of the Clinical Scales.

This study also focused on the issue whether the female subpopulation reported more (intense) traits than the male subpopulation. Regarding scales 1, 2, and 3, also known as the Neurotic Triad, both the female (in cluster 2 of the 4-cluster solution) and male (in cluster 1 of the 4-cluster solution) subpopulation scored significantly higher than their other-gender counterparts. Overall, the male subpopulation scored higher on scale 4, thereby partially confirming the second hypothesis of this study. Perhaps, seen as scale 4 incorporates past contact with the justice system, this finding sheds another light on the conclusion of Megargee et al. (1999). On the one hand, it could have been expected that the female and male subpopulation in this study would score equally high on scale 4. If female defendants are indeed more socially deviant compared to women with no history of contact with the justice system, it would result in more extensive criminal histories (resembling criminal histories of male defendants). This expectation was partially met: in all cluster profiles with clinical elevations, both the female and male subpopulation had elevated scores on scale 4. On the other hand, the issue whether incarcerated women might be more socially deviant than non-incarcerated women when compared to incarcerated versus non-incarcerated men could explain males' higher scores on scale 4: it 'takes more' for women to be arrested than it does for men, resulting in less contact with the justice system. The male subpopulation scored higher (but average) on scale 5. Regarding scales 6, 7, and 8, the female subpopulation scored significantly higher than the male subpopulation. This could mean that the female defendant population reports more intense feelings of guilt, insecurity, isolation and paranoid behavior than the male defendant population. No gender differences were found on Scales 9 and 0.

Several limitations of this study must be mentioned. First, in light of further

explanation, information concerning (the absence of) a diagnosis (Axis I, Axis II) would have

been desirable to check whether there was any resemblance with the MMPI-2 profiles. Second,

the decision on the number of clusters based on hierarchical cluster analysis remains the domain

of professional judgment. In principle, any number of clusters can be chosen as long as it

provides adequate information concerning the population one seeks to describe. This study served

as a replication of the Spaans et al. (2009) study in a different defendant population and extended

to a female subpopulation. For that reason, the statistical protocol adhered to by Spaans et al.

(2009) was followed. Third, information regarding the location of test assessment could point out

whether the results were subject to the mental state of the defendant at the time of test

assessment. For example, perhaps the defendants who reported severe personal/emotional distress

in the very disturbed clusters all resided in a House of Detention. The results could then be

greatly affected by their emotional state, which confounds the MMPI-2 profiles. One of the

present study's strengths was the large amount of participants/MMPI-2 questionnaires available.

It constituted a great part of all Dutch defendants awaiting trial in the period of 2006-2014, which

strengthens possibilities for external generalization.

The 4-6(-7)-8 profile, found in most of the clusters in this study,  reflects a general

outward hostility while experiencing recurring feelings of anxiety, worry and slight alienation;

this profile views the world as a survival zone. The '1-2-3-4-6-7-8(-0)' MMPI-2 profile emerged,

for all cluster solutions, in the 'very disturbed' cluster. Participants in these clusters report the

most psychological distress, even though it might be an exaggeration of symptoms in order to

receive help for personal/emotional distress. This clinical scale profile, including all the MMPI-2

symptom scales, fits the structure for a borderline personality structure (Nasiri, Abedi, Ebrahimi,

Ameli & Samouei, 2013). In a more general explanation, this profile concerns hostility, feelings

of insecurity, anxiety, worry, depression, antisocial tendencies, psychopathy, paranoia and a

33

general inability to trust others. This study implicates that the MMPI-2 in a general population of

Dutch pre-trial defendants, at best, can distinguish a 'normal' cluster, a disturbed cluster entailing

paranoid schizophrenic traits and symptoms, and a very disturbed cluster that features a

borderline personality structure. Moreover, these clusters have been found in two different

studies performed independently from each other, yielding near identical results. The type of

profiles one desires to find therefore seems to be subject to number of clusters chosen by the

researcher. The use of the MMPI-2 in the forensic pre-trial settings, therefore, can be questioned.

In favor of its use is the widespread acceptance of the MMPI-2 as an appropriate personality

questionnaire. It is often used as part of a test battery, in combination with other personality

questionnaires, or in combination with structured interviews and observations (Gordon, 2007).

From the results of this study, I recommend that the MMPI-2 has to remain part of a larger

assessment and should not be the only source of information when evaluating a defendant.

However, in defense of the MMPI-2, identifying defendants that experience severe

personal/emotional distress from defendants that do not is, on its own, a valuable distinction. This

acknowledgement has been stated by Grover (2011) as well. Moreover, it could even be argued

that three clusters for the general defendant population and two clusters for both female and male

defendants might actually accurately reflect this population. One of the present study's strengths

was the large amount of participants/MMPI-2 questionnaires available. It constituted a great part

of all Dutch defendants awaiting trial in the period of 2006-2014, which strengthens possibilities

for external generalization. To test this, information regarding Axis I and Axis II diagnoses of the

defendants is a requirement. Future research should focus on the 'fit' between MMPI-2 outcomes

and the clinical judgment of the psychologist.

## References

Aneshensel, C.S., Rutter, C.M., & Lachenbruch, P.A. (1991). Social structure, stress, and mental health: Competing conceptual and analytic models. *American Sociological Review, 56*(2), 166-178.

Butcher, J.N., Dahlstrom, W.G., Graham, J.R., Tellegen, A., & Kaemmer, B. (1989). *The Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring*. Minneapolis, MN: University of Minnesota Press.

Butcher, J.N., Graham, J.R., Ben-Porath, Y.S., Tellegen, A., Dahlstrom, W.G., & Kaemmer, B. (2001). *MMPI-2 (Minnesota Multiphasic Personality Inventory-2): Manual for administration, scoring, and interpretation, revised edition.* Minneapolis: University of Minnesota Press.

Cima, M., Merckelbach, H., Hollnack, S., Butt, C., Kremer, K., Schellbach-Matties, R., & Muris, P. (2003). The other side of malingering: Supernormality. *The Clinical Neuropsychologist, 17*(2), 235-243.

Cima, M., Van Oorsouw, K. (2013). The relationship between psychopathy and crime-related amnesia. *International Journal of Law and Psychiatry, 36*(1), 23-29.

Davis, K.M., & Archer, R.P. (2010). A critical review of objective personality inventories with sex offenders. *Journal of Clinical Psychology, 66,* 1254-1280. doi:10.1002/jclp.20722

Derksen, J., de Mey, H., Sloore, H. & Hellenbosch, H. (2006). *MMPI-2 Handleiding voor afname, scoring en interpretatie*. Nijmegen: PEN Tests Publisher.

Dienst Justitiële Inrichtingen. (2013, May). *Gevangeniswezen in getal 2008-2012*. Retrieved August 19, 2014, from http://www.dji.nl/Organisatie/Feiten-en-cijfers/index.aspx

Edvinsson, D., Bingefors, K., Lindström, E., & Lewander, T. (2010). ADHD-related

symptoms among adults in out-patient psychiatry and female prison inmates as compared with the general population. *Upsala Journal of Medical Sciences, 115,* 30-40.

Espelage, D.L., Cauffman, E., Broidy, L., Piquero, A.R., Mazerolle, P., & Steiner, H. (2003). A cluster-analytic investigation of MMPI profiles of serious male and female juvenile offenders. *Journal of the American Academy of Child & Adolescent Psychiatry, 42*, 770-777.

Fazel, S., & Danesh, J. (2002). Serious mental disorder in 23 000 prisoners: A systematic review of 62 surveys. *The Lancet, 359,* 545-550.

Formann, A.K. (1984). *Die Latent-Class-Analyse: Einführung in die Theorie und Anwendung*. Beltz: Weinheim.

Ginsberg, Y., Hirvikoski, T., & Lindefors, N. (2010). Attention Deficit Hyperactivity Disorder (ADHD) among longer-term prison inmates is a prevalent, persistent and disabling disorder. *BMC Psychiatry, 10*(112), 1-13.

Gordon, R.M. (2007). The powerful combination of the MMPI-2 and the Psychodynamic Diagnostic Manual. *Independent Practitioner*, 84–85.

Grover, B.L. (2011). The utility of MMPI-2 scores with a correctional population & convicted sex offenders. *Psychology, 2*(6), 638-642.

Hall, G.C.N., Graham, J.R., & Shepherd, J.B. (1991). Three methods of developing MMPI taxonomies of sexual offenders. *Journal of Personality Assessment, 56*, 2-13.

Hathaway, S.R., & McKinley, J.C. (1940). A multiphasic personality schedule (Minnesota): I. Construction of the schedule. *Journal of Psychology, 10*, 249-254.

Hathaway, S.R., & McKinley, J.C. (1942). A multiphasic personality schedule (Minnesota): III. The measurement of symptomatic depression. *Journal of Psychology, 14*, 73-84.

Hatters Friedman, S., Hall, R.C.W., Sorrentino, R.M. (2013). Commentary: Women, violence, and insanity. *Journal of the American Academy of Psychiatry and the Law, 41*, 523–528.

Heilbrun, K., & Heilbrun, A. (1995). Risk assessment with MMPI-2 in forensic evaluations. In Ben-Porath, Y.S., Graham, J.R., Hall, G.C.N., Hirschman, R.D., & Zaragoza, M.S. (Eds.), *Forensic applications of MMPI-2*. Thousand Oaks, CA: Sage.

Karno, M., Hough, R.L., Burnam, M.A., Escobar, J.I., Timbers, D.M., Santana, F., & Boyd, J.H. (1987). Lifetime prevalence of specific psychiatric disorders among Mexican Americans and non-Hispanic whites in Los Angeles. *Archives of General Psychiatry, 44*, 695-701.

Maguire, K., & Pastore, A.L. (Eds.) (1995). *Bureau of Justice Statistics: Sourcebook of criminal justice statistics--1994.* Washington, DC: U.S. Department of Justice.

McGilloway, A., Hall, R.E., Lee, T., and Bhui, K.S. (2010). A systematic review of personality disorder, race and ethnicity: Prevalence, aetiology and treatment. *BMC Psychiatry, 10*(33), 1-14.

McKee, G.R., Shea, S.J., Mogy, R.B., & Holden, C.E. (2001). MMPI-2 Profiles of Filicidal, Mariticidal, and Homicidal Women. *Journal of Clinical Psychology, 57*(3), 367-374.

Megargee, E.I. (1997). Using the Megargee MMPI-based classification system with the MMPI-2s of female prison inmates. *Psychological Assessment, 9*, 75-82.

Megargee, E. I., & Bonn, Jr., M. J., Meyer, Jr., J., & Sink, F. (1979). *Classifying criminal offenders: A new system based on the MMPI.* Beverly Hills, CA: Sage.

Megargee, E.I., Carbonell, J.L., Bohn, M.J., Jr, & Sliger, G.L. (2001). *Classifying Criminal Offenders with the MMPI-2: The Megargee System.* Minneapolis, MN: University of Minnesota Press.

Megargee, E.I., Mercer, S.J., & Carbonell, J.L. (1999). MMPI-2 with male and female state and federal prison inmates. *Psychological Assessment, 11*, 177-185.

Nasiri, H., Abedi, A., Ebrahimi, A., Ameli, S.S., & Samouei, R. (2013). Personality profile of women affected with borderline personality disorder. *Materia Socio Medica, 1*, 60-63.

Nieberding, R., Gacono, C.B., Pirie, M., Bannatyne, L.A., Viglione, D.J., Cooper, B., … Frackowiak, M. (2003). MMPI-2 based classification of forensic psychiatric outpatients: An explanatory cluster analytic study. *Journal of Clinical Psychology, 59*, 907-920.

Robins, L.N., HeIzer, J.E., Weissman, M.M., Orvaschel, H., Gruenberg, E., Burke, J.D., & Darrel, A.R. (1984). Lifetime prevalence of specific psychiatric disorders in three sites. *Archives of General Psychiatry, 41*, 949-958.

Sloore, H., Derksen, J., De Mey, H., & Hellenbosch, G. (1996). The Flemish/Dutch version of the MMPI-2: Development and adaptation of the inventory for Belgium and The Netherlands. In: J.N. Butcher (Ed.). *International adaptations of the MMPI-2: Research and clinical applications*. Minneapolis: University of Minnesota Press.

Spaans, M., Barendregt, M., Muller, E., de Beurs, E., Nijman, H., & Rinne, T. (2009). MMPI profiles of males accused of severe crimes: A cluster analysis. *Psychology, Crime & Law, 15*(5), 441-450.

Van Oorsouw, K., & Merckelbach, H. (2011). Detecting malingered memory problems in the civil and criminal arena. *Legal and Criminal Psychology, 15*(1), 97-114.

Verma, J.P. (2013). *Data analysis in management with SPSS software*. India: Springer. doi: 10.1007/978-81-322-0786-3_10

Vinkers, D.J., Barendregt, M., De Beurs, E., Hoek, H.W., & Rinne, T. (2011). Etnische verschillen tussen pro Justitia gerapporteerde verdachten. *Tijdschrift voor Psychiatrie,*

*53*(11), 801-811.

Weiner, I.B., Freedheim, D.K., & Graham, J.R. (2003). *Handbook of psychology, assessment psychology*. Hoboken, NJ: Wiley & Sons.