SERIES: 'MYTHS AND HYPES IN I/O PSYCHOLOGY'

The Undesired Popularity of Typologies and other 'Jung'[1]

Patrick Vermeren[2]

This article appeared in the Dutch scientific Magazine 'Gedrag & Organisatie' *(Behaviour & Organizations)*

(**Patrick Vermeren is an author, HR consultant for PerCo, and chairman of the non-profit agency vzw Evidence Based HRM. Correspondence address: Kleine Meylstraat 4C, B-2550 Kontich; e-mail: vermeren.consulent@telenet.be)

ABSTRACT. Typologies are still very popular in organisations and companies, and in commercial consultancy firms (providers, especially of training and coaching). The assumptions of psychiatrist Carl Gustav Jung, in particular, seem to be in favour again, based on the many new tests that have been introduced recently. This article discusses the problems with the foundations of Jungian theory (psychoanalysis, paranormal beliefs, and metaphysical archetypes), Jungian tests, and their test reliability, or lack of it. Two tests based on Jungian typologies are discussed: the MBTI® (Myers-Briggs Type Indicator®) and Insight Discovery®. I demonstrate the widespread use of typologies: Various companies, renowned institutes, such as business schools and university departments, as well as government authorities and municipal services use them and even ask for them. I discuss the reasons for the current state of affairs and recommend making an effort to turn the tide in favour of evidence-based practice.

1. Introduction

In 2006, Academia Press published my book (available only in Dutch) *De HR-Ballon. 10 Populaire Praktijken Doorprikt*. In the epilogue 'Getuigenis van een spijtoptant' (Testimony of an individual who regrets a chosen option), I look back in embarrassment at my time as a trainer at a Belgian bank (Vermeren, 2006, p.195). Naivety, gullibility, and various forms of bias often result in internal trainers and external commercial suppliers being sucked along in

[1] With thanks to Reinout de Vries, Tom Speelman, Johan Braeckman, and Maarten Boudry for their comments on previous versions. Thanks to psychometrician Danny Rouckhout for the text on the problems with ipsative tests. I also wish to thank the anonymous reviewer 2 and Bernard Nijstad for their valuable comments.

[2] Patrick Vermeren is an author, HR consultant for PerCo, and chairman of the non-profit agency vzw Evidence Based HRM. Correspondence address: Kleine Meylstraat 4C, B-2550 Kontich; e-mail: vermeren.consulent@telenet.be

the propagation of popular, but scientifically debatable, models and tests, and I was no exception. Scientifically debatable models and tests such as the MBTI (Myers-Briggs Type Indicator), transactional analysis, Situational Leadership® II, the Thomas-Kilmann conflict instrument, Maslow's pyramid of needs, the educational theories of Kolb, and the 'grief cycle' of Elisabeth Kübler-Ross are still extremely popular, despite unfavourable assessments or even the abandonment of the theory by the authors themselves (Lilienfeld, Lynn, Ruscio, & Beyerstein, 2010, p.62). Sadly, little appears to have changed in the meantime. In particular, typologies are extremely popular amongst commercial suppliers and clients.

Typologies are as old as the road to Rome. In ancient times, Hippocrates and Galenus divided people into types according to the mixture of 'the four bodily fluids': the sanguine temperament (warm blooded), the choleric temperament (hot tempered, due to too much yellow bile), the melancholic temperament (due to too much black bile), and the phlegmatic temperament (due to too much phlegm). We still come across these ideas in our daily use of language (warm-blooded, hot-headed, and phlegmatic). Furthermore, the psychoanalytical range of thought hasn't entirely disappeared—certainly not in organisations and companies. Although not so true of Freud's ideas, Jung's archetypes are not only unusually popular, they are also undergoing a genuine re-emergence. However, many people don't know the origin of these archetypes and types and whether or not they have a sound foundation.

This article is a critical evaluation of Jung's theories and the tests based upon them. First, I discuss the theoretical foundations of Jungian theory, and then I consider the problems with Jung's typologies and the tests that are based on those typologies. Next, I consider the spread and popularity of typology-based tests, and I conclude with a discussion of some of their harmful effects and some lines of thought for resolving this issue.

2. Jung's theory and its foundation

In 1921, Carl Gustav Jung (1875-1961), a Swiss doctor who was interested in people with mental disorders, described a number of 'types' that he claimed were based on the speculative work of William James (1842-1910; Bair, Fontijn, Nieuwkoop, & Visser, 2004, p.292 and 328). Jung's typology draws a distinction between three dimensions. First, he argued that each of us is born with a dominant attitude or style (extroversion or introversion). Second, people were said to have 'irrational' information processing preferences, namely experience (sensory perception) versus intuition. Third, he postulated that people have 'rational' preferences, reflected in the dichotomy of thinking versus feeling. Jung claimed that these rational and irrational preferences (which he called cognitive or mental functions) formed a hierarchy within the personality. According to Jung, of the four 'functions' (thinking, feeling, sensing, and intuition), one would be dominant and a second would be what he called an auxiliary function. Furthermore, the auxiliary functions develop with age. Jung arrived at a combination of these three dimensions, which he described in his book *Psychological Types* (1921). On a side note, it is difficult to tell if Jung envisaged two,

four, eight, or even sixteen types, because of his ambiguous writing style, filled with neologisms and complex terms (which he said were designed to frighten off lunatics).

The theoretical foundations of Jung's ideas originate from psychoanalysis as well as belief in the paranormal, telepathy, and mythological thought. Psychoanalysis is largely based on the idea that psychological problems are caused by forgotten or repressed conflicts, experiences, and desires from childhood. The originator of course was Sigmund Freud, who used the term *psychoanalysis* for the first time in 1896. He introduced talk therapy on the basis of his hypotheses and provided the following definition: "Psychoanalysis is the name (1) of a procedure for the investigation of mental processes which are almost inaccessible in any other way; (2) of a method (based on that investigation) for the treatment of neurotic disorders; and (3) of a collection of psychological information obtained along those lines, which is gradually being accumulated into a new scientific discipline" (Freud, 1923, p.211).

Initially, Jung was an adherent of Freud's theories. The official version was that Jung turned away from Freud because the latter too often attributed psychological problems to repressed sexual trauma. Jung developed his own 'school of thought' and drew some of his ideas from mythology. Initially, Jung developed ideas about a collective evolutionary unconsciousness (1912), using as a base Ernst Haeckel's speculations (1834-1919) that during pregnancy foetuses develop psychologically and physically according to a number of stages which are analogous to millions of years of evolution. Jung thought he could describe this collective unconsciousness in 'archetypical symbols' and in Chinese Taoist alchemy (trying to convert base metals into gold). Jung developed the concepts of anima and animus, analogous to the Chinese concepts of Yin and Yang. He developed the hypothesis that the 'false self' has to disappear to make way for the 'true self.' Men can use their feminine side (anima) and women can use their masculine side (animus) to discover their true self via dreams and active imagination.

Jung also developed his own psychoanalytical ideas. For example, he thought up the terms *individuation* (a personal development process that establishes a connection between the 'ego' and the 'self') and the *shadow* (a form of archetype that contains all negative characteristics that the individual wishes to deny, including our animal instincts). He also believed that dreams served to restore psychological balance.

The second foundation of Jung's ideas was his faith in the paranormal and telepathy. For example, he borrowed the term *synchronicity* from the Lamarckian biologist Paul Kammerer (1880-1926), who concluded on the basis of his diary that strange concurrences of circumstances could not be based on coincidence. Jung defined synchronicity as 'meaningful coincidences' in which mental processes coincide in time with 'phenomena in the world of perception.' The concept of synchronicity led Jung to claim that all patients over the age of 35 could be helped through the knowledge of archetypes that are found in the collective unconsciousness. These archetypes are not a result of the physical world, but exist separately in a 'parallel universe' (metaphysical – the third foundation). However, according

to Jung, every human brain has access to this world. Psychological disorders arise because people are beleaguered by spirits from the metaphysical world of the collective unconsciousness. Access to the collective unconsciousness would provide the solution to psychological disorders.

Jung's faith in the paranormal and the metaphysical archetypes, in combination with psychoanalysis, prompted him to develop his now renowned 'Types.' These types constitute the foundation for a number of popular tests, such as MBTI®, TDI®, JTI, Insights Discovery®, and Golden Personality Type Profiler™ (see below). What these tests have in common is that they operate on the assumption that all people can be divided into a limited number of types that differ from each other qualitatively. Some of these tests are based on Jung's three original dimensions (e.g., Insights Discovery); most of the tests are based on four dimensions (e.g., MBTI, TDI, and JTI); and the Golden Personality Type Profiler is based on Jung's original three dimensions plus two extra dimensions.

3. What's wrong with the Jungian theory and tests?

3.1 Problem 1: Unsound theoretical foundations

The first major problem is that there is no empirical evidence for the key concepts of psychoanalysis, paranormal phenomena, and mythology. Concepts from psychoanalysis have been refuted piece by piece: Memory research has shown that 'unconscious repression' does not exist (e.g., Loftus, 1994a and 1994b); the interpretation of dreams does not yield any workable hypotheses (either according to Freudian or Jungian interpretation; e.g., Lavie & Hobson, 1986; Hobson, Pace-Schott, & Stickgold, 2000), schizophrenia is not caused by 'regression of the libido' (as claimed by Freud), and autism is not caused by insensitive, 'refrigerator mothers.' Autism and schizophrenia both have a strong genetic component and are currently regarded as developmental disorders of the brain. For example, on the basis of twin research, several research groups have calculated that the heritability of schizophrenia is 80% to 84% (Cardno et al., 1999; Kendler, Myers, Potter, & Opalesky, 2009). Similarly, there is no empirical support for other ideas from psychoanalysis such as penis envy (i.e., all girls are said to envy boys because of their penis) or the Oedipus complex (i.e., whereby every boy between the age of three and five secretly dreams of having sex with his mother and killing his father). A more extensive discussion of the problems of psychoanalysis as a theory and as a therapy is outside the scope of this article, but let me conclude here by referring readers to the study led by epidemiologist Yolba Smit, which resulted in the discontinuation of reimbursement for psychoanalytical therapy by health insurers in the Netherlands (Smit et al., 2010; 2012), or to more extensive reviews (e.g., Buekens, 2006).

Furthermore, Jung's ideas stemming from mythology and metaphysics also lack empirical evidence. No one has ever provided any evidence for the existence of synchronicity, and when Jung defended himself, he often made use of fallacies such as, "… because statistics are possible only if there are exceptions" (Adler et al., 1973 C.G. Jung Letters, vol. 2, p.246). Jung was a notorious believer in alchemy, astrology, spiritism, telepathy, telekinesis,

clairvoyance, and extrasensory perception. Historical sources have shown that he was influenced by William James, who, amongst other things, believed in communications with spirits via media in séances. Jung maintained that there are people who sense things 'intuitively,' for example that 'a yellow car will come around the street corner.' Jung also cited examples of this 'from his practice': A gold (scarab) beetle that flew against the window as a patient was relating her dream which featured a beetle, showed, according to Jung, that there has to be a non-coincidental link between the mental world and the phenomena such as this one from the physical world (Jung, 1960, p.142). So far, however, no one has been able to demonstrate paranormal gifts or extrasensory perception under controlled conditions, despite a reward of one million US dollars offered by James Randi several years ago. Recently, that sum was supplemented by one million euros by the Belgian non-profit organisation Skepp.

As described earlier, Jung postulated that of the four functions (thinking, feeling, sensing, and intuition) one was dominant and another one was auxiliary. The difference between the dominant function and the auxiliary function should be reflected in a test as a higher score for the dominant function and a lower score for the auxiliary function. In the MBTI this is expressed in the so-called JP index (Judgment-Perception Index), which was designed to determine a person's dominant function. However, the existence of dominant and auxiliary functions has never been confirmed in research, neither in studies by the Myers themselves (Myers & Myers, 1980) nor by others (e.g., McCrae & Costa, 1989). Moreover, test results have not shown that auxiliary functions develop with age.

If anyone is still in any doubt after this consideration of the main 'theoretical foundations,' let me repeat that Jung's typologies have never been proven empirically. The excuse that they cannot be tested using current scientific methods is not very convincing. Jung did not conduct any scientific studies, and he based his theories mainly on his own observations and anecdotal accounts during a period in which many people described him as psychologically sick. The major reason why tests based on Jung's 'theory' are almost always given an unfavourable assessment is because they are based on *an unsound theoretical foundation* (e.g., the assessment of the MBTI by the COTAN).

Jung—just like Freud—never made any secret of the fact that he did not follow the path of academic science:

"Anyone who wants to know the human psyche will learn next to nothing from experimental psychology. He would be better advised to abandon exact science, put away his scholar's gown, bid farewell to his study, and wander with human heart through the world." (Jung's *New Paths in Psychology, Collected Works*, London, 1916).

3.2 Problem 2: Type is at odds with biological variation

Tests that are based on Jung's ideas generally divide people into a series of discreet types. However, the reasoning behind typology contains a major fallacy, namely the assumption of

dichotomy and bipolarity. Practitioners operate on the principle that the scales are *discontinuous*, *dichotomous*, or *bimodal*. This would mean that the population can be divided into two groups per scale, with a 'gap' in the middle of each distribution. Compare this to the idea, for example, that the male population consisted of two groups: men between 1.40 and 1.60m tall and men between 1.80 to 2m tall, with hardly any men between 1.60 and 1.80m tall. However, human characteristics are distributed normally, whether it is a question of height, muscle power, intelligence, or personality traits. Differences in personality are therefore more gradual or 'fluid.' Modern personality psychology favours the trait approach–these traits are presented in continuous scales or dimensions instead of in dichotomies. The most accepted and scientifically established model is, of course, the Big Five or the Five Factor Model (FFM–five major trait domains). The most well known and researched test is the NEO-PI-R by Costa and McCrae (1995). Both the five major domains and the thirty underlying facets (six per domain) show a normal distribution. Whichever scale one chooses, the mutual combinations of the many facets of our personality produce a vast potential for variation between personalities! A new domain with a number of facets may emerge because now there are data that appear to show a sixth domain called 'honesty' (HEXACO model, Lee & Ashton, 2004).

Typologies do not take into account the gradual differences in personality and the enormous variation of human characteristics and their possible combinations: For example, MBTI states that people can be divided up into sixteen types; LIFO® uses four types; and Enneagram uses nine types. According to a typology, a person definitely belongs either in one category or the other. In other words, one category excludes the other. People are either extrovert or introvert; however, the bulk of the population is neither extrovert nor introvert, but lies somewhere in between (referred to as ambivert). The use of ipsative tests (forced choice) emphasises the dichotomy (see below). Any form of typology reasoning is problematic for two main reasons: (1) the enormous variation in personalities as described above, which is an effect of evolutionary influences (e.g., random mutations or the arbitrary mixture of genes as a result of sexual selection), other biological influences (e.g., hormonal influences during pregnancy or viruses), and developmental processes; and (2) the influence of environmental or contextual factors that cause people to react differently in different situations (Barkow, Cosmides, & Tooby, 1992; Moscowitz & Zuroff, 2004).

And so, variations between people are caused mainly by evolutionary processes and are certainly not limited to four, eight, or sixteen types. There is little room for doubt because the theory of evolution is considered to be one of the best-founded scientific theories (Coyne, 2010; Dawkins, 2009; Dennet 1996; Williams, 1996). Evidence for this theory has been supplied throughout history from many branches of science, such as DNA research, genetics, anthropology, and archaeology. Therefore, some people call the theory of evolution a theorem, which indicates that it is a theory or hypothesis that has been indisputably proven.

Hans van Gossum (1) summarises evolution in four principles, which together offer an explanation for the personality differences between people.

1. There is variation between different individuals. You only have to look at a school class photograph to realise that people are (outwardly) different in several respects.

2. There are always more children born than the number who grow to become adults. Nature provides limited resources, therefore a lot of plant seeds and young animals perish.

3. The individual who adapts best to the environment has more chance of survival. These adaptations are made possible by genetic variation. There are two known mechanisms that ensure a certain variation gets the upper hand. The first is coincidence: When genetic features randomly spread across a population, it is referred to as 'genetic drift.' The second mechanism is natural selection.

4. Beneficial features (adapted or 'fit' to the environment) are passed on by means of procreation. Only those who can successfully procreate will spread the beneficial features, which means a population can adapt to changing circumstances.

Thus, the origin of differences between people can be explained parsimoniously by the theory of evolution, but not on the basis of Jung's 'parallel metaphysical world.'

3.3 Problem 3: fictitious and incorrectly used scales

Jungian theory operates on the existence of three dichotomies, yet all three have been subjected to scientific criticism. First, there is the dichotomy of experience versus intuition. Intuition, as described by Jung, arose from his faith in the paranormal, but, as I have already said, there is no evidence for this. Intuition is a concept deployed in modern psychology, but with a different definition: It is the whole of implicit knowledge acquired by multiple experiences in a regular and therefore predictable environment and the opportunity to learn these regularities by lengthy practice (Kahneman, 2011, p.252). This holds up well for professions such as fire fighting, medicine, and nursing, but not for professions such as financial investment advice, political science, and psychotherapy. Similarly, the feeling versus thinking scale is based on an untenable dichotomy. Research in both clinical psychology and neurobiology (e.g., Damasio et al., 2001) has shown that a distinction cannot be made between emotions and thoughts. In fact, they are linked to each other indivisibly in neural networks in the human brain. Clinical psychologists (especially those trained in cognitive behavioural therapy) have firmly adopted that point of view. Anger, for example, is always related to the same sort of thoughts, namely thoughts that involve a command or prohibition. This often finds expression in thoughts involving the words *must* or *not allowed* or *that's not possible* because a particular desire or objective is under threat. Fear is always related to thoughts that express negative effects: 'That dog will bite me,' 'My partner will be angry,' 'The client won't like that,' or 'The dentist is going to hurt me.' In terms of human behaviour, as is the case with animals, fear almost always leads to some form of avoidance,

flight, fight, or freeze. Obviously, some people are less stable emotionally than others, but that has nothing to do with intellectual capacity. It is scientifically untenable to portray people who easily become afraid as less capable in 'thinking.' In the FFM or the Big Five, emotional stability is represented as a dimension (highly unstable to highly stable). However, the question of how prudently someone can think is a completely independent dimension.

Finally, the way certain tests based on Jung's ideas (such as the MBTI) deal with the scale of extroversion versus introversion is problematic. These tests present this scale as a dichotomy. However, Jung himself argued that there was no such thing as a person who was only extrovert or introvert, and that these were factors or dimensions (this is also what contemporary research in this field has shown–see above). Jung said that anyone who was only extrovert or introvert should be "admitted to an asylum." And so, it is the developers of the MBTI and other Jungian typology tests who have introduced this dichotomy and formulated introvert versus extrovert as a type antithesis. A proposition such as 'introverts draw their energy from within themselves, whilst extroverts draw energy from others' cannot be tested from a scientific point of view and is in contradiction with other scientific disciplines such as physics.

3.4 Problem 4: The tests are unreliable, artificially reliable, or unresearched.

There is a problem in putting Jung's archetypal theory into operation, and the problem lies in the tests themselves. Some of these tests have been researched for their psychometric quality, but the results have proved extremely problematic. With the MBTI, for example, there is up to a 60% chance of a person being classified under a completely different type after just four weeks–the test-retest reliability is unacceptably low (see below under MBTI).

Without going into an analysis of all the problems again, I will briefly address the ipsative nature of these tests (2) in a little more depth. Ipsative scoring is a system whereby the respondent actually divides up a set number of points (the constant value) over (usually) a number of scales that are included in the test. Therefore, the sum of the different item scores will be equal for each respondent. The result is an arrangement of the scales on the basis of their importance to the respondent (intrapersonal). This also explains the term ipsative: Indeed, the term *ipsative* comes from *ipse* = self. There are various ways of obtaining ipsative scores, such as preference scores or comparison in pairs, but the most commonly used form is the forced choice answer form. For example, the set of items in the MBTI always consists of two items which represent the poles of a bipolar scale and one of the two has to be chosen.

The biggest problem with ipsative scores lies in the artificial reliability of the tests. Amongst other things, factor analysis is necessary to demonstrate construct validity and starts from an intercorrelation matrix, for example, of the different scales. However, intercorrelation matrices based on ipsative scores have one very peculiar feature: That is, the average intercorrelation of all correlations from the matrix is always equal to -1/(k-1), in which k

stands for the number of scales in the test. Therefore, irrespective of the items in the scales, the mathematical calculation of how the scales intercorrelate can be done in advance. For example, Meade (2004) collected the test data of 2,895 candidates for the job of sales assistant within a big supermarket group in the USA. The scores for the eight different scales were collected and reproduced in a matrix in both ipsative and normative ways. The predicted average intercorrelation is therefore equal to -1/(8-1) or -0.14. The real average intercorrelation in this matrix was indeed -0.14 (p. 543). On the basis of this rule, it is also clear that the number of scales will increase as the average intercorrelation falls. With fifteen scales, the average intercorrelation is equal to -1/(15-1) or -0.07. As the number of scales rises, the individual correlation values will gradually decrease, therefore a high correlation in the matrix will rarely be encountered. Such matrices are often incorrectly interpreted as evidence for a set of independent variables, certainly whenever the reader is not aware of the ipsative nature of the scales. In other words, the independency of the scales is purely artificial.

The intercorrelation matrix of a series of ipsative scales is an artefact of which there is no meaningful interpretation. If researchers carry out a factor analysis on such artificial matrices, they will see even more peculiar effects which have been well described by other authors (Baron, 1996; Bartram, 1996; Clemans, 1966; Dunlap & Cornwall, 1994; Hicks, 1970; Johnson, Wood, & Blinkhorn, 1988; Tenopyr, 1988). These effects are of a statistically technical nature (such as, production of artificial bipolar factors; sometimes commonalities equal to 1 and the absence of specific variance; the lower the scales, the greater the artificial distortion; and such like). Dunlap and Cornwall (1994, p.123) concluded: “We are left with recommending against the use of principal component, principal factoring, or maximum likelihood factor analysis with ipsative measures. The separation of artifactual bipolar factors induced by ipsativity from any true underlying relationship will be difficult at best, and not worth the danger of a largely incorrect interpretation.” Tenopyr (1988, p.750) states: “And so, it is clear that the internal consistency reliabilities of forced-choice scales are not only interdependent but may also augment each other enough to give the scale user a false sense of confidence in construct interpretation,” and “… a subject for major concern in view of the widespread use of forced-choice inventories in vocational guidance, diagnosis, selection, and research.”

Basing one’s conclusions on ipsative test results can have major effects: The results often lead people to completely different conclusions than when using normative scales. To illustrate this, Meade established that recruitment decisions based on ipsative scales were totally different from those based on normative scales. Although some studies appear to show that ipsative scales can sometimes be appropriate in vocational preference tests, for example, they are not appropriate for measuring personality dimensions or for selection (Meade, 2004, p.548). Since the MBTI uses ipsative scales, both the internal reliability of the scales and the independence of the scales are unreliable because they are artificial.

4. Jung is in da house

On the Internet, typologies (and typology tests) based on the ideas of Carl Gustav Jung are extremely popular. The web site of Type Association Benelux (downloaded on 12 January 2012 from www.type-association.org) includes the MBTI (Myers-Briggs Type Indicator – 16 types), and it also includes the following instruments which refer explicitly to Jung's personality types: MTR-I (Management Team Roles Indicator); TDI (Type Dynamics Indicator); JTI (Jungian Type Indicator); and Insights Discovery, an instrument that is emerging rapidly in Belgium. A recent newcomer, which is also based on Jung, is the GPTP (Golden Personality Type Profiler). Below I briefly discuss only two of these tests for want of any independent publication about the reliability and validity of the others. Yet, in fact, a discussion has little point because most scholars and scientific philosophers regard a model, typology, or test without a sound theoretical foundation as futile.

4.1 The ever popular MBTI

The MBTI was developed in 1942 by Isabel Briggs Myers and Katherine Cook Briggs (her mother). The MBTI (Briggs & Myers, 1987) is an ipsative (forced choice) questionnaire which, according to proponents, indicates the course of a preference on four bipolar dimensions (Hicks, 1970): extroversion-introversion (E-I), sensing-intuition (S-N), thinking-feeling (T-F), and judging-perceiving (J-P). The MBTI theory specifies that there are three levels at which a typological distinction could be made: (1) qualitative differences between the opposite preferences (listed for that purpose); (2) statistical interactions between preferences on the basis of external criteria (such as performance); and (3) the difference between a dominant function and an auxiliary function.

Briggs Myers and Cook Briggs claimed that Jung differentiated on the basis of four instead of three dichotomies. In *Psychological Types,* Jung clearly referred to three distinct dichotomies: extrovert versus introvert attitudes and two opposite pairs of functions: rational (judging) functions (thinking-feeling) and the irrational (perceiving) functions (sensing-intuition). The MBTI developers added the 'judging versus perceiving' (judging-perceiving) dimension. This dimension should indicate whether someone prefers to plan his or her life according to plans and structure (J) or in a more flexible manner (P). And so, this dimension differs with regard to content from Jung's written contentions.

Other basic principles of the MBTI have emerged that are contrary to Jung's theory. I have already discussed the problematic dichotomy of extroversion versus introversion, which was criticised by Jung himself. The fact that the authors regard each type as fundamentally equivalent and positive (Myers & Myers, 1980) is untenable because people with pronounced tendencies often have conflicts or problems. For example, dominant and aggressive leaders cause a lot of adverse effects and are often designated in the subject literature as *abusive supervisors* (e.g., Tepper, 2000). McCrae and Costa (1989) argue that the Big Five dimension of neuroticism (the measure of emotional stability) on its own provides evidence that not all 'types' can be regarded favourably.

McCrae and Costa (1989) conducted research into the soundness of the MBTI because they were looking for a good personality test for their research at the National Institute on Aging. Whilst the Big Five theory emerged as solid, they also researched whether or not the MBTI was a good way of putting Jung's theory into operation. And so, (once again) they examined whether or not people really could be divided up into sixteen types, whether there really were dichotomous preferences, whether there were interactions between the preferences (55 possible interactions in all), and whether or not people developed their auxiliary functions alongside their naturally dominant functions as they got older, as Jung claimed. McCrae and Costa (1989, p.32) found no confirmation for any of the claims of the MBTI developers summarised above: "There was no evidence that preferences formed true dichotomies, the 16 types did not appear to be qualitatively distinct, because analyses of their joint effects on personality dimensions showed that only 1 of 55 interactions was significant, and only in women, and, contrary to hypothesis, the theoretically dominant function was no more clearly preferred than the auxiliary. The Jungian prediction that opposing functions should be developed in later life was not confirmed using the MBTI." McCrae and Costa also argued: "Weighing the evidence to date, the MBTI does not seem to be a promising instrument for measuring Jung's types, those who embrace Jung's theory should probably avoid the MBTI." The authors also argue that either Jung's theory is faulty or the fault lies with the way in which it is put into operation, namely the MBTI.

The third level that MBTI theoreticians describe (dominant versus auxiliary functions) should, for example, be found easily in the test results: Dominant functions must get a higher (preference) score in these tests than auxiliary functions (Myers & McCaulley, 1985, p.58). This could not be confirmed in any study at all, not even in their own research, although they didn't seem to realise that if these interactions could be found amongst only half of the types (Myers & McCaulley, 1985, p.60), this yields the same result as what could be expected on the basis of pure coincidence.

The reliability of the test is also problematic. Independent research has shown that 47% to 60% of MBTI test participants arrive at a different result after five weeks. Howes and Carskadon established in a 1979 study that the test-retest reliability was very weak: After a test-retest interval of a mere five weeks, 50% of people showed up under a different type. In 1983, McCarley and Carskadon replicated these results. In a study carried out by the National Research Council (NRC, 1991), it emerged that "from all studies examined, only 24 to 61% of the participants showed stability in type." All together, 15 psychologists worked for the NRC on the MBTI test and investigated 11 relevant studies with test-retest data. Their findings showed that 39% to 76% of participants were given a different type in a repeat of the test administered no more than three to five weeks later. The NRC calculated a median of 60% uncertainty regarding the allocated type. To be precise, this would mean that over half of the people change their personality type each month!

Similarly weak test-retest reliability results have been shown by scholars such as Bess and Harvey (2002), Fleenor (2001), and Mastrangelo (2001). Isabel Myers herself reported in 1998 that 35% of the test persons had different type-scores after a test-retest interval of only four weeks. As shown, that percentage is much higher in independent studies. Other researchers, too, argue that the Big Five tests are clearly superior to the MBTI (e.g., Furnham, 1995; Pittenger, 1993 and 2005). Although correlations have sometimes been found between the scales of Big Five tests and MBTI, that does not provide support for the proposition that the theoretical foundation of the MBTI is sound. And, as discussed previously, the internal reliability of the instrument is obviously artificial as in all ipsative tests.

The Dutch Committee on Tests and Testing (COTAN) of the Dutch Association of Psychologists (NIP) gives the following assessment of the MBTI in its *Documentation of Tests and Test Research*:

- Basic principles of the test construction: insufficient; unsound theoretical foundation
- Quality of the test material: good
- Quality of the manual: good
- Standards: insufficient; standards too small and standards not representative and/or the representation cannot be assessed
- Reliability: good
- Concept validity: insufficient; not enough research
- Criterion validity: insufficient (no research).

In short, most academics generally agree that the MBTI is not a sound instrument and that there are much better tests available.

4.2 Insights Discovery

The Insights Discovery instrument (referred to as 'evaluator' in the quotation below) has been heralded as follows: "The Insights Discovery model is based on the extensive research of Swiss psychologist Carl Jung and the subsequent work of Jolande Jacobi, one of his leading students." And: "The evaluator has been tested and updated to measure the quality of the 100 word pairs and weaker ones replaced with stronger ones. They've also been tested for reliability and validity in huge numbers to help gauge how robust the word pairs used in the evaluator are," (downloaded on 12 January 2013 from

www.insights.com/2119/validating-the-insights-discovery-model.html). The following can be read about this instrument elsewhere: "There is still a lot of unrelenting scientific interest in the work of Jung." (Frank Sample, 2005, p.5; www.christelclear.nl/files/Frank%20Sample%20-%20Basis-Man-Pers.ontw.%20NL.pdf; downloaded on 12 January 2013).

Apart from the incorrect claim that Carl Gustav Jung was a psychologist (he graduated in medicine in 1902 and specialised in psychiatry), there are also problems with the claim that the theory is sound and that the test is valid and reliable. Research in the APA database on 12 January 2013 did not yield even a single article on Insights Discovery (0 hits). There is also no entry in the COTAN documentation (Documentation of tests and test research in the Netherlands–NIP or Dutch Association of Psychologists, 12 January 2013); nor does the Buros Institute have even a single review (consulted on 12 January 2013). This is a classic: People are making strong claims about the scientific status of an instrument because they reckon that the vast majority of people will not test it or are not capable of assessing it properly.

4.3 The other Jungian tests

There is a lack of evidence for the soundness of the other tests. Apart from the journals devoted to Jungian typology and psychoanalytical publications, such as the *Journal of Psychological Type*, there is not a single reference to be found about most of these tests in sound, peer-reviewed journals. A consultation of the APA database on 12 January 2013 yielded the following results: for TDI, only a summary definition of the test with no peer review; for the rest, 0 hits; for MTR-I, only a summary definition of the test which has not been subjected to peer review; for the rest, 0 hits.

5. How widespread are these typologies?

Individual companies don't like to reveal which tests their HR or training departments use. Psychologist and journalist Dominique Haijtema (2008) discovered this in her journalistic research. However, she was able to determine that the Dutch tax office, KLM, Shell, and Hay Group were all using at least one of the typology tests mentioned previously. Haijtema purports that it is easy to ascertain which companies use or have used a particular test by visiting the web sites of the commercial test providers. This method may not be reliable in the case of consulting companies that act as test providers because these companies are suspected of self-interest; hence, other indicators must be used to demonstrate how widespread and popular they are.

A simple Internet search provides an indication of how widespread typologies are amongst companies and which ones are using them. Obviously, other type theories and instruments, such as the Enneagram (nine types), the LIFO (Life Orientations–only four types), and the Belbin® team roles (nine team role types), are still popular. Yet the test with the most hits on the Internet is the MBTI. The four letters *MBTI* on their own yielded about 4,500,000 hits

via the Google search engine on 30 December 2012. About 16,000 pages are from Belgium and 48,500 are from the Netherlands. Below is a partial list of major institutions that use MBTI, sometimes in their lesson packet:

- Business Schools: The Vlerick Business School from Ghent (downloaded on 30 December 2012 from www.vlerick.be) uses MBTI in its 'Middle Management Programme'; the RMS or Rotterdam School of Management (Erasmus University Rotterdam) also uses it in their management training courses (downloaded on 31 December 2012 from www.rsm.nl/about-rsm/news/detail/1511-eeod-tip-sheet-getting-the-most-out-of-mbti/). The web site of the IMD international school from Switzerland (downloaded on 30 December 2012 from www.imd.org) shows that the MBTI is used, for example, in their programme *Mobilizing People, Leadership, Coaching Teamwork*, and the INSEAD from France (www.insead.edu) uses it in several programmes. IMD and INSEAD explicitly recruit top people for their 'executive development programme.' A search on their various web sites (downloaded on 30 December 2012) reveals that RMD, IMD, and INSEAD also use Belbin team roles.

- Universities: The web sites of Ghent University (Faculty of Economics and Business; downloaded on 30 December 2012) and Twente University (Faculty of Management and Business, downloaded on 30 December 2012) show these institutions use it in their training courses.

- Government Agencies: The Government Agencies Staff (AgO) of the Flemish Community (AgO training brochure, p.60) states that it uses MBTI "to get to know itself and its preferences better"; the cities of Ghent (public specifications for "the provision of training for managers and leaders for the employees of the City of Ghent," 2008; p.31) and Antwerp (training catalogue for employees, p.206, also with Belbin, p.201) are a few examples from Belgium alone.

- Research driven companies: AstraZeneca and Novartis are two companies that are obliged to carry out medical research before their medicines are allowed on the market. Thus, they should be familiar with scholarly research; yet according to www.insights.com, both companies use the Insights Discovery. At AstraZeneca, according to Insights, the test is used "to improve team engagement and effectiveness" (downloaded on 12 January 2012 from www.insights.com/45/customer-case-studies.html).

The problem is that these organisations are helping to perpetuate the impression that the MBTI is a sound test that is respected by scholars. This is not a responsible scholarly approach. As for service providers, such as consultants and business schools, this certainly seems to be the result of financial considerations and laziness–it is easy to earn money with a popular test. Universities are apparently houses with so many rooms that the faculty members don't even bother to ask their psychology professors for advice.

6. Why have these Jungian typologies survived?

It is obvious that providers of tests who are fully aware of the problems, but do not point them out, are the biggest culprits. One publisher of tests recently tried to evade responsibility by suggesting to me, "Yes, but there's a demand for those tests, so we supply them."

Providers such as HR consultants and trainers are certainly responsible in part for the widespread use of these tests. However, a number of these providers act in good faith because they are untrained in major fields such as behavioural biology, evolutionary psychology, personality psychology, and labour psychology. There is no doubt that some consciously opt for the easy money that can be earned by offering the most popular models, but a good number of them act simply out of ignorance and have fallen for (false) lines of authoritative reasoning and manipulation techniques (Cialdini, 2009). However, judging by the vast range of models, typologies, and tests available from providers of training and coaching, they do not seem to be making any effort to distinguish between sound and unsound theories. They may not be capable of distinguishing between good and bad scholarship and don't take the trouble to steep themselves in the scholarly literature of the fields mentioned above. Research in Belgium (Segers, Vloeberghs, De Prins, & Henderickx, 2009) and the Netherlands (Groen, Sanders, & Van Riemsdijk, 2006; Sanders, Van Riemsdijk, & Groen, 2008) has confirmed that HR professionals have very little training in psychology, let alone training to assess the soundness of certain research articles. This shows that there is a very low level of knowledge of academic findings on HR concepts and instruments for HR professionals. This applies to both external providers (consultants, trainers, and coaches) and internal HR staff and line managers.

Apart from popular claims such as *'*millions of people use it and they can't all be wrong*,'* which are discussed extensively in the subject literature (e.g., Bardone & Magnani, 2010; Cialdini, 2009; Goodwin, 1998; Walton, 2000), four commonly heard rationalisations for the use of Jungian typology are as follows:

> 'It doesn't matter if it's scientific or not, as long as you get people talking about it.'
>
> 'Sometimes typologies can be useful to help you communicate reality and make it somewhat comprehensible.'
>
> 'In my experience people get a lot out of it,' or 'I've had good experiences with it.'
>
> 'Science often contradicts itself; maybe they'll be saying ten years from now that XXX is a good test.'

These arguments ignore a number of important scholarly conclusions. For example, people soon think that (vague, positive, or ambiguous) psychological statements apply to them– certainly if people calling themselves experts are the ones making such statements (the so-

called Barnum effect). The fact that people recognise themselves or are satisfied with the results does not constitute evidence for the reliability of the knowledge. A second issue is what you can actually impart to people if you give them a fundamentally flawed theoretical framework and/or the test results are very unreliable. As far as 'good experiences' are concerned, it is precisely because our own experiences often yield unreliable knowledge (due to the many forms of *bias* or intuitive prejudices) that we have set up the scholarly method. Obviously, this method is far from being infallible, but it always yields much more reliable knowledge than subjective experiences or 'hearsay.' People who appeal to their own experience as proof are thereby actually showing that they are not familiar with the scholarly method and its benefits, or have no idea of the problematic effects related to excessive trust in one's own subjective experiences (Kahneman, 2011, p.53). The claim that science repeatedly contradicts itself is an easy ploy to kill off debate, but the reality is that in most sciences there is a body of progressive insight. Mechanisms are often refined or new ones discovered, but the main points remain intact (e.g., the findings from the theory of evolution, see Coyne, 2009; Dennett, 1995). Furthermore, the common assumption that typologies do offer some help and can be useful in certain contexts has never yet been proven in any way whatsoever. The databases of APA, ABI-Inform, and Business Source Elite do not contain any empirical evidence for this proposition.

The clients (the business community and organisations) using these tests may also be guilty by association in some cases, or may at least be accused of neglect. Even in sectors such as chemistry and pharmacy where evidence-based work practices should be taken for granted because scientific research is at the core of their business, the HR department often does not work in an evidence-based manner. This is probably due to the selection criteria (some HR departments have few organisational or other psychologists in their midst). HR is still treated harshly in many companies and few HR managers are found in company management committees. Senior managers from the line often determine the HR course and even the models that they want deployed in the company. Moreover, they ignore the advice of labour psychologists who warn them about the dubious nature of a theory. In addition, the internal training department frequently has the autonomy to make these decisions, whether evidence based or not. Nevertheless, big organisations should be perfectly capable of setting aside the financial resources to give at least one person within the organisation access to the major databases (e.g., APA, PsycArticles, ABI Inform, Elsevier, and Wiley) and giving that person enough time to study scholarly literature on HR and management. Companies should make a policy choice to work in an evidence-based manner where possible. Moreover, many scientists suggest using the term 'science-based,' presumably due to the erosion of the concept of evidence-based policy.

The academic world could also be blamed for neglect. The fact that these typologies are still so popular is, as far as I can see, also due to the lack of a unifying theory, so the field is extremely splintered and the different domains seem to refute each other with incompatible theories and poorly defined concepts. The social sciences are split up into

several domains and, unlike the natural sciences (e.g., physics and chemistry), do not have a tradition of scientific integration and remain isolated from the findings from other domains. The Jungian theory can then be seen by some as just one of the competing models that is no better or worse than any other.

Furthermore, it appears that little effort is being made to publicly—outside the 'ivory tower'—refute flawed theories and myths, with a few modest exceptions. That is a pity because this widespread use of typology tests is bad for psychology as a discipline. It certainly feeds the public scepticism towards psychology as a scientific discipline (see also Lilienfeld, 2012). The internal focus on scientific publications rather than popular publications (i.e., 'publish or perish') and the time that it takes for an individual to struggle against well oiled marketing machines (to say nothing of intimidation techniques such as threats of legal action) certainly play a part here.

Sadly, even less can be expected from psychologists who are not connected to a university. Academic research at least has a number of self-cleansing mechanisms, as demonstrated by the current debate raging in universities about the effects of publish or perish. But things are worse amongst the professional association of psychologists in Belgium (personal correspondence 15 December 2010) and non-psychologists in the HR sphere of work: At present, there simply *isn't any* intrinsic form of quality control. Observing the doggedness with which some psychologists-practitioners stubbornly regard the Jungian types as sound, one realises how weak the effect is of warnings against confirmation bias so often heard during psychology classes and how difficult it is to 'sell' the scientific method.

The employees who undergo the training courses or tests are (generally speaking) not psychologists themselves. All people are gullible and impressionable by nature towards (seemingly) authoritative arguments (e.g., 'Jung, psychiatrist and father of modern psychology'; 'scientifically tested'; 'our HR department uses this'; 'business schools use this'; etc.) and the 'bandwagon' effect of social evidence (e.g., 'So many people and companies can't be wrong!'). This allows people to spare themselves any effort and gives a false sense of security. People are often not aware of the fact that they are continually susceptible to these sorts of errors of judgment (*bias*) due to the division of the brain into functional (network) modules (Kurzban, 2010). Daniel Kahneman (2011) simplifies the module division of the human brain into the categories 'System I' and 'System II,' in which System I stands for rapid impressions, gullibility, positive feelings, and the tendency to see patterns or causal links everywhere (which are often not there at all). System II stands for slow, considered, and critical reasoning which requires a lot of effortful deliberation, a heartrending (critical) frame of mind, and conscious attention to mental effort (Kahneman, 2011, p.76). The 'lazy' System I often gets the upper hand, so it's not surprising that employees are not well informed about the theoretical foundations and the way typologies are put into operation. Moreover, our education system does not teach us to think critically. People also fall into the trap of the promise of instant success: Marketing materials for a type model usually

purport the model offers a relatively simple explanation for the complexity of human personality and behaviour, that people will be able to assess other people more accurately, and that people can better gear their communication to others so they are in a better position to achieve their own objectives.

In my view, typologies are a special sort of poison—a dangerous 'meme'. A major problem that results from the unreliable human brain is the tendency to put people, animals, and things into categories (categorisation). 'Type thinking' or putting people into boxes fits in perfectly with this tendency and even encourages this division. Categorisation helps us to simplify the world, but it also leads to problems such as prejudices, in-group versus out-group thinking, and racism. Furthermore, despite all warnings (e.g., not to use a test for selection purposes) and privacy legislation, people fail en masse on these points. I regularly see CVs in which people quote their 'MBTI type' and I also sometimes go into organisations where the profile of the entire team is displayed on the wall. Finally, it seems easy to imagine that lots of employees will refrain from criticising the choices of their manager or the HR or training department out of fear of a negative impact on their careers.

7. Discussion

It can be concluded from the information presented above that tests based on Jung's theories are problematic. This isn't only a problem in the practice of test usage, there are other possible detrimental effects. Think of the harmful effects for employees, the harmful effects for the image of psychology (people will see it as a pseudoscience), the harmful effects for the organisations who use these tests, the harmful effects for those who offer tests in good faith, and the harmful effect of stimulating type-thinking.

Scholarly literature makes no mention of research into harmful effects as a result of the use of such problematic theories and tests. Nevertheless, there are strong indications of this in the related field of psychotherapy. For example, in recent years, researchers have conducted studies in the field of mental health care that show substantial damage can arise. It has emerged that certain forms of therapy yield harmful effects, such as false accusations against parents (who are thought to be a cause of psychological problems), worsening of behavioural problems, and heightened risk for posttraumatic stress (Lilienfeld, 2007). Moreover, research into harmful effects has only just started in standard clinical psychology, so evidence of even more adverse effects can be expected (Lilienfeld, 2007). A lot of psychoanalytical ideas (from the many and sometimes contradictory schools of thought), such as a death wish, the cause of all potential psychological problems must be sought in the first three years of one's life and in one's sexual desires, the Oedipus complex, and the now heavily contested view that 'uncaring mothers' are a possible cause of autism (see, for example, the documentary *Le Mur* by Sophie Robert, which was initially banned due to pressure from the psychoanalytical lobby), have not been extinguished amongst scores of non-academic psychologists, let alone amongst practitioners who often lack any psychological training. When the 'origin statements' (where the problems come from) of

psychoanalytical theories and the tests in training sessions and individual coaching which are based on them are outlined, then it is not unthinkable that some people will lay the blame for their psychological problems at the feet of their parents. People can adapt their self-image in the wrong way on the basis of bad tests, which can led to faulty career decisions (e.g., 'As an *Introverted Feeler*, I'm not suitable for a career as a manager'), with the accompanying financial effects and even psychological wounds. As in clinical psychology, there is a pressing need for research into harmful side effects.

Psychology as a scientific domain has everything to lose by perpetuating the psychoanalytical range of thought of Freud, Lacan, and Jung. Obviously, scientifically oriented psychologists look upon all this with sorrow, but, even worse, this is pure pseudoscience in the eyes of academics from subjects like (behavioural) biology. The forecast by Darwin that psychology would eventually have to align with biology is still a long way off. Only a few domains, such as behavioural genetics and evolutionary psychology, have gone down that road. Nevertheless, as Trivers (2012, p.306) argues, psychology has to continue to build on more hard sciences. Physics is supported by mathematics, chemistry is supported by physics; biology is supported by chemistry, and psychology should be supported by biology. His outlook is shared wholeheartedly by other biologists, such as Richard Dawkins, and also by psychologists, such as David Buss, Steven Pinker, Leda Cosmides, and John Tooby. They see the brain as an organ that has been 'designed' by evolution to process inputs from the environment. The brain is a complete system of cells (physics) that communicate with each other via chemical (chemistry) and electrical signals. Just like other animals, humans have at their disposal a number of biological dispositions, such as sexual jealousy and a drive for status amongst males (biology). Psychology cannot ignore these underlying processes. Human beings are characterised by a universal human nature which can be found in all cultures of the world. Cultures certainly display more similarities than differences, and the differences are probably already the effect of other environmental inputs (e.g., 'evoked culture'—as a result of what food is present and to what extent; Barkow, Cosmides, & Tooby, 1992). The need to share with other scientific domains is quite pressing.

The common lack of a metatheory in the social sciences, and in psychology in particular, is coming under increasingly sharp criticism. For example, Robert Trivers, recognised for his contributions to evolutionary biology, states: "Like economists, psychologists were going to create their field out of itself; learning theory, social psychology, psychoanalysis—essentially competing guesses about what was important in human development, none with any foundation," (Trivers, 2012, p.315). Scientific philosophers and psychologists themselves have regularly taken shots at this. For example, Hans Eysenck (1997, p. 1224) stated: "It is suggested that the scientific status of psychology is put in danger by the lack of paradigms in many of its fields, and by the failure to achieve unification, psychology is breaking up into many different disciplines." To further illustrate, I quote Jesse Marczyk: "I find your lack of theory (and replications) disturbing … / … without theory, all you have is a grab bag of

findings, some positive, some negative, and no idea what to do with them or how they are to be understood," (consulted on 14 May 2013 on www.popsych.org).

Another, often underexposed, problem related to these typologies is that providers of nonscientific or proven faulty models, theories, and tests are competing with sincere, evidence-based providers. They do this by using a number of tried and tested manipulation techniques. Apart from the claims of scientific reliability and validity, as outlined above, these providers also make frequent use of genuine, false, or assumed authoritative arguments, such as those made by self-proclaimed experts (Caroll, 2000). Here are two such statements from the Internet: "The renowned Swiss psychiatrist and psychologist Carl Gustav Jung," "TDI is continuing to work on the typologies of Carl Jung (1875-1961), one of the fathers of modern psychology," (downloaded on 31 December 2012 from www.amelior.be). Jung wasn't a psychologist and he was merely one of the fathers of psychoanalysis. Furthermore, they also make use of the bandwagon-effect or *popular* argument (e.g., 'millions of people have taken this test'), social evidence via so-called *testimonials* (i.e., public testimonies and recommendations), and so on.

The use of problematic theoretical concepts is a potential time bomb within organisations. Many people who take typology tests based on Jung's theory have no idea of the foundations of these typologies. They are ignorant of the fact that the paranormal belief in one of the domains (intuition) has been completely refuted. For example, would they know that Jung obtained his doctorate in research into the occult and that he continued to believe in the paranormal all his life? When employees find out that the Jungian archetypes and the types based on them are intrinsically occult and 'come from a parallel universe,' this can do a lot of damage to the credibility of HR and training and development. If such concepts and tests were used for selection and recruitment, this could lead to financial damage claims from applicants who feel they have been taken in. Indeed, according to some people, it is only a matter of time before such legal action is seen in our part of the world (Western Europe). Finally, I find it remarkable how few people seem to take offence at the fact that someone is being dishonest towards participants, namely by providing an incorrect frame of reference and withholding essential information about the problems with the theories and test qualities.

The solution that I see is to work in an evidence-based manner, and perhaps there is a ray of light on the horizon. Recently, there have been several initiatives to provide a counterweight, including the formation of CEBMA (Centre for Evidence Based Management – www.cebma.org) and the non-profit organisation VZW Evidence Based HRM (www.evidencebasedhrm.be). Furthermore, the Belgian VOV-Lerend Netwerk (formerly called the 'Vereniging voor Opleidings- en Vormingsverantwoordelijken'), and the Vigor Innovation Group (www.vigorinnovation.com–an initiative of the Ghent University and the Catholic University of Leuven), have jointly set up a learning network 'Evidence-Based HR' (3). Academics, who embrace the methods of science and critical thought, must provide

leadership. They can help their 'drifting fellow human beings' escape from faulty ideas such as typologies and other things by:

- Framing their research within a unifying theory, which must be sought in the evolution of man as a biological being. Psychology needs to develop a broader view of science and take a look over the wall at other disciplines by (1) tracing the convergence or contradiction of their research results and (2) inviting scientists from these disciplines (e.g., biologists) to commentate on their work. Once again, psychology has to finally look to biology for more support, as Darwin already suggested (1859, p.488): "In the distant future I see open fields for far more important researches. Psychology will be based on a new foundation, that of the necessary acquirement of each mental power and capacity by gradation." In my opinion, evolutionary psychology provides this unifying theory.

- Being a beacon in the dark by setting up an organ with a mission similar to that of The Cochrane Collaboration relative to the medical world. This organ could not only publish systematic reviews, but also give a clear division of journals on the basis of research quality.

- By taking action publicly if nonsense appears in magazines or newspapers, and certainly if HR professionals are the writer's target group.

- By conducting independent (!) research (which might or might not be for money) at the request of 'the field' into certain subjects that are considered to be relevant in daily practice.

At the very least, the information role is a function of the universities in society, mainly because universities are financed by taxpayers' money. In a democracy, diligence is necessary to ensure that research funds are used as effectively as possible, even if it is only to make sure that health care remains affordable, through the elimination of non-evidence-based practices and the advancement of proven techniques.

Practice box

What do the results mean in practice?

- This literature review shows that there are strong arguments for abandoning typology thinking in the domain of personality in favour of dimensions, namely the Five Factor Model (FFM) or the Big Five of personality traits. The overwhelming majority of academic scientists have already trodden this path, but the business community is limping behind and arguably even heading down a dated and disproved road.
- There are two other important reasons for abandoning typologies. First, there are indications in clinical psychology that some practices and the application of certain theories could cause harmful effects, and the same could apply to the use of (unsound) tests. Second, this article sounds a precautionary warning for the potential deontological and judicial effects.
- Furthermore, it is important to be aware of the fact that people are all susceptible to various forms of bias which are a hindrance to well-founded choices: falling for attractive simplicity (categorisation) and positivism and the impact of (supposedly) authoritative arguments. A conscious effort is required to resist these biases and look critically at models and tests.
- By applying some rules of thumb, people are in a better position to assess certain theories or tests for soundness of evidence. Practitioners are advised to look at various scientific domains for areas of convergence (or the lack of it) of the findings and to focus on systematic reviews. Moreover, they can base their efforts on a number of low-threshold publications and web sites from recent initiatives such as www.evidencebasedhrm.be or www.cebma.org.
- The author believes that academics can play a special role in the dissemination of evidence- based information and practice.

Notes

1. Hans van Gossum is a researcher from The Research Foundation of Flanders, and a member of the Evolutionary Ecology study group of Antwerp University. Source: special contribution on evolution in the EOS journal.
2. An extensive discussion of the problems with ipsative tests can be found in the keyword register on the web site http://evidencebasedhrm.be. This is an excerpt from the literal text that was written by psychometrician Danny Rouckhout.

3. See http://vov.be/wat-doet-vov/lerende-netwerken/netwerk-evidence-based-hr/.

References

Adler, G., Jaffe, A., & Hull, R.F.C. (1973). *C.G. Jung Letters Vol. 2,* 1951-1961. Princeton, NJ: Princeton University Press.

Bair, D., Fontijn, B., Nieuwkoop, P., & Visser, W. (2004). *Jung: Een biografie*. De Bezige Bij.

Bardone, E., & Magnani, L (2010). The appeal of gossiping fallacies and its eco-logical roots. *Pragmatics & Cognition, 18*, 365-396.

Barkow, J., Cosmides, L., & Tooby, J. (eds) (1992). *The adapted mind: Evolutionary psychology and the generation of culture.* New York: Oxford University Press.

Baron, H. (1996). Strengths and limitations of ipsative measurement. *Journal of Occupational and Organizational Psychology, 69*, 49–56.

Bartram, D. (1996). The relationship between ipsatized and normative measures of personality. *Journal of Occupational and Organizational Psychology, 69*, 25–39.

Bess, T.L., & Harvey, R.J. (2001, April). *Bimodal score distributions and the MBTI: Fact or artifact?* Paper presented at the Annual Conference of the Society for Industrial and Organizational Psychology. San Diego.

Briggs, K., & Myers, I. (1987). *Myers-Briggs Type Indicator Form G.* Palo Alto, CA: Consulting Psychologist Press.

Buekens, F. (2006) *Freuds vergissing. De illusies van de psychoanalyse*. Leuven: UMCO / Uitgeverij Van Halewyck.

Cardno, A.G., Marshall, E.J., Coid, B., Macdonald, A.M., Ribchester, T.R., Davies, N.J., Venturi, P., Jones, L.A., Lewis, S.W., Sham, P.C., Gottesman, I.I., Farmer, A.E., McGuffin, P., Reveley, A.M., & Murray, R.M. (1999). Heritability estimates for psychotic disorders: The Maudsley twin psychosis series. *Archives of General Psychiatry, 56*, 162 - 168.

Carroll, R.T. (2000). *Becoming a critical thinker - A guide for the new millennium*. Boston: Pearson Custom Publishing.

Cialdini, R. B. (2009). *Invloed. De zes geheimen van het overtuigen*. Den Haag: Academic Service.

Clemans, W. V. (1966). An analytical and empirical examination of some properties of ipsative measures. *Psychometric Monographs, 14*.

Costa, P. T., Jr., & McCrae, R. R. (1995) Domains and facets: Hierarchical personality assessment using the Revised NEO Personality Inventory. *Journal of Personality Assessment, 64*, 21-50.

Coyne, J. (2009). *Why evolution is true.* Penguin Books.

Damasio, A.R. (2001). Emotion and the human brain. In A.R. Damasio, A. Harrington, J. Kagan, B.S. McEwen, H. Moss, & R. Shaikh (Eds.), *Unity of knowledge: The convergence of natural and human science*. *Annals of the New York Academy of Sciences, 935*, 101-106. New York: New York Academy of Sciences.

Darwin, C.R. (1859) *On the origin of species, by means of natural selection*. London: John Murray.

Dawkins, R. (2009) *The greatest show on Earth: The evidence for evolution*. London: Transworld Publishers.

Dennett, D. (1995). *Darwin's dangerous idea*. London: Penguin Press.

Dunlap, W. P., & Cornwell, J. M. (1994). Factor analysis of ipsative measures. *Multivariate Behavioral Research, 29*, 115–126.

Eysenck, H.J. (1997) Personality and experimental psychology: The unification of psychology and the possibility of a paradigm. *Journal of Personality and Social Psychology, 73*, 1224-1237.

Fleenor, J. W. (2001). Review of the Myers-Briggs Type Indicator Form M. In B.S. Blake, , & J.C. Impara, (Eds.) *The fourteenth mental measurements yearbook*. (pp. 816 – 818), Lincoln, NE: Buros Institute of Mental Measurements, The University of Nebraska Press.

Freud, S. (1923/1976). «Psychoanalyse» und «Libidotheorie». In S. Freud *Gesammelte Werke* 13. (pp. 209–233). Frankfurt am Main: S. Fischer Verlag.

Furnham, A. (1995). The big five versus the big four: The relationship between the Myers-Briggs Type Indicator (MBTI) and NEO-PI five factor model of personality. *Personality and individual Differences,* 21, 303-307.

Goodwin, J. (1998). Forms of authority and the real Ad Verecundiam. *Argumentation, 12*, 267–280.

Groen, B., Sanders, K.,& van Riemsdijk, M. (2006). De kloof tussen theorie en praktijk. Een onderzoek naar de kennis van HRM'ers over arbeids- en organisatiepsychologie. *Tijdschrift voor HRM, 9*, 33 - 52.

Hajtema, D. (2008). Wie bent u? Honderden dure testen brengen persoonlijkheid in kaart, maar de kwaliteit is dubieus. *Het Financieel Dagblad.* 6 september 2008, FD persoonlijk, 6 -9.

Hicks, L. E. (1970). Some properties of ipsative, normative, and forced choice normative measures. *Psychological Bulletin*, *74*, 167-184.

Hobson, J. A., Pace-Schott, E. F., & Stickgold R. (2000). Dreaming and the brain: Toward a cognitive neuroscience of conscious states. *Behavioral and Brain Sciences, 23*, 793–842.

Howes, R.J., & Carskadon, T.G. (1979). Test-retest reliabilities of the Myers-Briggs Type Indicator as a function of mood changes. *Research in Psychological Type, 2*, 67-72.

Johnson, C. E., Wood, R., & Blinkhorn, S. F. (1988). Spuriouser and spuriouser: The use of ipsative personality tests. *Journal of Occupational Psychology, 61*, 153–162.

Jung. C.G. (1912). *The psychology of the unconscious. (vertaling van Wandlungen und Symbole der Libido)* London: Kegan Paul Trench Trubner.

Jung. C.G. (1921). *Psychologische typen*. Zurich: Rascher Verlag. Translation in G. Adler & R.F.C. Hull (Eds. & Trans), *Collected Works of C.G. Jung* (1957-1979), 'Psychological Types' (volume 6, 1971). Princeton, NJ: Princeton University Press.

Jung, C.G. (1950). *Foreword. The I Ching or Book of Changes.* (Translation) Princeton, NJ: Princeton University Press.

Jung, C.G. (1957 – 1979). Psychology and Alchemy (volume 12, 1968) and Alchemical Studies (volume 13, 1968). In G. Adler & R.F.C. Hull (Eds. & Trans.), *Collected Works of C.G. Jung*. Princeton, NJ: Princeton University Press.

Jung, C.G. (1957-1979). The Structure and Dynamics of the Psyche (volume 8, 1960). In G. Adler & R.F.C. Hull (Eds. & Trans.), *Collected Works of C.G. Jung*. Princeton, NJ: Princeton University Press.

Jung, C.G. (1964). *Man and His Symbols*. New York: Doubleday.

Kahneman, D. (2011). *Thinking, fast and slow.* New York: Farrar, Straus and Giroux.

Kendler, K.S., Myers, J., Potter, J., & Opalesky, J. (2009). A web-based study of personality, psychopathology and substance use in twin, other relative and relationship pairs. *Twin Research and Human Genetics, 12*, 137-141.

Kurzban, R. (2010). *Why everybody (else) is a hypocrite: Evolution and the modular mind.* Woodstock (UK): Princeton University Press.

Lavie, P., & Hobson, J. A. (1986). Origin of dreams: Anticipation of modern theories in the philosophy and physiology of the eighteenth and nineteenth centuries. *Psychological Bulletin, 100*, 229–240.

Lee, K., & Ashton, M. C. (2004). Psychometric properties of the HEXACO personality inventory. *Multivariate Behavioral Research, 39*, 329–358.

Lilienfeld, S.O. (2007). Psychological Treatments That Cause Harm. *Perspectives on Psychological Science, 2*, 53.

Lilienfeld, S.O. (2012). Public skepticism of psychology. Why many people perceive the study of human behavior as unscientific. *American Psychologist, 67*, 111–129.

Lilienfeld, S.O., Lynn, S.J., Ruscio, J., & Beyerstein, B.L. (2010). *50 Great myths of popular psychology. Shattering widespread misconceptions about human behavior*. Chisester UK; Wiley-Blackwell.

Loftus, E. (1994a). *The myth of repressed memory.* New York: St. Martin's.

Loftus, E. (1994b). The repressed memory controversy. *American psychologist, 49*, 443–445.

Marczyk, J. (2013). I find your lack of theory (and replications) disturbing. www.psychologytoday.com. Blog post in May 2013.

Mastrangelo, P. M. (2001). Review of the Myers-Briggs Type Indicator, Form M. In B.S. Blake,, & J.C. Impara, (Eds.) *The fourteenth mental measurements yearbook*. (pp. 818 – 820), Lincoln, NE: Buros Institute of Mental Measurements, The University of Nebraska Press.

McCarley, N.G., & Carskadon, T.G. (1983) Test-retest reliabilities of scales and subscales of the Myers-Briggs Type Indicator and of criteria for clinical interpretive hypotheses involving them. *Research in Psychological Type*, *6*, 24-26.

McCrae, R. R., & Costa, P. T., Jr. (1989). Reinterpreting the Myers-Briggs Type Indicator from the perspective of the Five-Factor Model of personality. *Journal of Personality, 57*, 17-40.

Meade, A.W. (2004). Psychometric problems and issues involved with creating and using ipsative measures for selection. *Journal of Occupational and Organizational Psychology, 77*, 531–552.

Moskowitz, D.S., & Zuroff, D.C. (2004). Flux, pulse, and spin: Dynamic additions to the personality lexicon. *Journal of Personality and Social Psychology, 86*, 880-893.

Myers, I.B., & McCaulley, M.HK (1985). *Manual. A guide to the development and use of the Myers-Briggs Type Indicator*. Palo Alto, CA: Consulting Psychologists Press.

Myers, I.B., McCaulley, M.HK, Quenk, N.L., & Hammer, A.L. (1998). *Manual: A guide to the development and use of the Myers-Briggs Type Indicator*. Palo Alto, CA: Consulting Psychologists Press.

Myers, I. B., & Myers, P. B. (1980). *Gifts differing*. Palo Alto, CA: Consulting Psychologists Press.

National Research Council; Druckman, D., Bjork, R.A. (Eds.) (1991). *In the mind's eye: Enhancing human performance.* Commission on Behavioral and Social Sciences and Education. Washington, D.C.: National Academy Press.

Pittenger, D. J. (1993). Measuring the MBTI ... and coming up short. *Journal of Career Planning and Employment, 54*, 48-53.

Pittenger, D.J. (2005). Cautionary comments regarding the Myers-Briggs Type Indicator. *Consulting Psychology Journal: Practice and Research, 57*, 210.

Sanders, K., van Riemsdijk, M., & Groen, B. (2008). The gap between research and practice: A replication study on the HR professionals' beliefs about effective human resource practices. *The International Journal of Human Resource Management, 19*, 1976-1988.

Segers, J ., Vloeberghs, D., De Prins, P., & Henderickx, E. (2009). Niets is zo praktisch als een goede theorie: Wie weet dit in HRM?. *Tijdschrift voor HRM, 2*, 7-28.

Smit, Y., Huibers, M., Ioannidis, J., van Dyck R.,van Tilburg, W., & Arntz, A. (2010). The effectiveness of psychoanalysis - a systematic review of the literature. (studie uitgevoerd voor het College van Zorgverzekeraars – www.cvz.nl).

Smit, Y., Huibers, M., Ioannidis, J., van Dyck R., van Tilburg, W., & Arntz, A. (2012). The effectiveness of long-term psychoanalytic psychotherapy--a meta-analysis of randomized controlled trials. *Clinical Psychology Review*, *32*, 81-92.

Tenopyr, M.L. (1988) Artifactual reliability of forced-choice scales. *Journal of Applied Psychology, 73*, 749-751.

Tepper, B. (2000). Consequences of abusive supervision. *Academy of Management Journal*, *43*, 178-190.

Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J. Barkow, L. Cosmides, & J. Tooby, (Eds), *The adapted mind: Evolutionary psychology and the generation of culture.* New York: Oxford University Press.

Trivers, R. (2002). *Natural selection and social theory. Selected papers of Robert Trivers*. New York: Oxford University press.

Trivers, R. (2012). *The folly of fools. The logic of deceit and self-deception in human life*. New York: Basic books.

Vermeren, P. (2006). *De hr-ballon. 10 populaire praktijken doorprikt*. Gent: Academia Press.

Walton, D. (2000). Evaluating appeals to popular opinion. *Inquiry: Critical Thinking Across the Disciplines, 20*, 33-45.

Williams G.C. (1996) *Plan and purpose in nature.* London: Weidenfeld & Nicolson.