



THE THEMATIC APPERCEPTION TEST IN THE LIGHT OF RESEARCH. THE POSSIBILITY OF USING THIS TEST IN FORENSIC ASSESSMENTS*

Katarzyna STEMPLEWSKA-ŻAKOWICZ

Faculty of Diagnostic Methods and Applied Psychology, Chodkowska University, Warszawa, Poland

Abstract

As the 1990s drew to a close, projective tests became the focus of discussion and heated debate, while new studies led experts to question their validity and practical utility. This probably explains why interest in projective techniques has recently begun to decline. The main charges laid at the door of projective techniques relate to their lack of validity and reliability, as well as doubts over whether psychopathology can be discerned in projected material. Drawing from these allegations, the author offers practical recommendations for applying projective techniques, and, while highlighting the need to use procedures that possess standards and proven validity and reliability, she draws attention to the need to attend training sessions devoted to the application of specific methods. She cautions against adopting the erroneous assumption that instances of sexual and physical abuse can be uncovered using projective techniques, while also drawing attention to the fact that, no matter what the research aim may be, these methods should be applied in a variety of ways. She discusses the legitimacy of the above recommendations by taking the Thematic Apperception Test as an example and relating it to several of the most popular systems for interpreting this test, i.e. those of Murray, Bellak, Westen and Cramer.

Key words

Projective methods; The statistical value of projective techniques; TAT; Systems of interpretation for TAT.

Received 27 December 2012; accepted 20 March 2013

1. Criticism of projective techniques

A significant proportion of projective diagnostic techniques appeared in the first half of the 20th century. Over the decades that followed, scholarly writing on the most familiar of these, such as the Rorschach and Thematic Apperception Tests, filled scores of academic titles every year. However, in the 1990s and at the turn of the new millennium something began to change. Projective techniques became the focus of discussion and heated debate, while new researches led experts to question their validity and practical utility

[16, 18, 23, 25, 28, 30, 80, cf. 64]. It was probably this criticism based on the conclusions of new studies that triggered the recent decline in interest in projective techniques. A search of the ECSCO databases for publications containing the term “Thematic Apperception Test” in their titles turned up 628 items published in peer-reviewed scholarly articles for the period 1935–2012, but limiting the search to the last 10 years only produced 33 publications (most of which were in one periodical – the *Journal of Personality Assessment*). Another sign of the new times is the fact that in the 2003 edition of his very popular handbook on clinical assessment, Gary Groth Marnat [20] chose not to include the chapter on projective drawing due to the rising number of empirical reports attesting to the doubtful validity of the indicators that were based on it¹.

* The paper was presented during the “Conditions of Using Projection Methods in Psychological Diagnosis for Court Purposes” Conference, which was held on 16–17 November 2012 in Lublin. It made use of small passages from other works by the same author [62, 63, 64].

¹ He was actually the author of one of these studies [21].

The main charges critics [30, 80] have aimed at projective techniques are in fact their lack of validity and reliability². Instead of these, impression-based interpretations are used that are not related to any standards, because as a rule such standards do not actually exist. For a long time, the objection has also been raised that psychopathology is being discerned in projected material and moreover, that the greater the quantity of projected material a given person “produces”, the more disturbances are effectively being “identified” in him. Bernard [52], in principle an advocate rather than one of the critics of projective techniques, cautioned against drawing such an erroneous inference. Due to the aforementioned weaknesses, projective techniques lack incremental validity, i.e. there is no observable improvement in the validity of a complete assessment after additional data from a projective technique have been made available for it [15]. Since it is clear that projective techniques cannot – in view of the aforementioned flaws – lead to valid conclusions, their application in psychological assessment should, in the opinion of their critics, be regarded as questionable. So what is to be done?

2. Practical recommendations for applying projective techniques

As far as the tendency to overestimate psychopathology is concerned, it is very possible that it is “in the eye of the beholder” and it would be prudent to regard this as one of the errors of informal diagnostic inference that Meehl [42] cautioned against. The remedy advised for this is for the diagnostician to extract herself from sole reliance on her experience or clinical intuition, and to base her diagnostic judgements on the confirmation – or non-confirmation – of the presence of “hard” indicators of psychopathology. Only, what indicators may we regard as “hard”? By way of response, Garb, Wood, Lilienfeld and Nezworski [16] formulated three conditions which an indicator should meet in order for it to form the basis of a responsible ruling on psychopathology. These conditions are as follows:

1. the indicator must be clearly linked to a given disorder, trait or syndrome;
2. the data proving this should be attained in methodologically sound studies;
3. the results should be independently replicated.

² And in some cases, the lack of research on this question, proving the careless attitude the authors of this technique displayed with regard to the basic requirements set for diagnostic tools [cf. 57, 65].

Therefore, it is not enough that a link of some nature be demonstrated in a single study. As the above authors [16] state, many very positive results relating to one or another projective indicator have already been attained. This should come as no surprise given the large number of studies on projective techniques conducted to date. However, those indicators that have been independently confirmed in the studies of another team are worthy of being taken into account in practice when their validity is not in doubt. It must be stated at this juncture that unfortunately, in Poland very many projective indicators are applied (in various techniques) that do not possess an empirical grounding of this kind. These indicators should be withdrawn forthwith. Some of these (regarding projective stories in the Thematic Apperception Test, TAT) will be pointed out in later sections of this article.

Various works by critical authors also contain pointers and recommendations for diagnosticians employing projective techniques. The author of this article gathered these together and presented them [62, 63] in the form of four recommendations, as follows:

1. use procedures (or systems of interpretation) whose validity and reliability has been adequately confirmed in studies and for which standards have been developed;
2. undergo training in the application of a given procedure before you start applying it in practice;
3. do not use projective techniques to uncover instances of sexual or physical abuse;
4. apply projective techniques in a variety of ways dependent on the research aim.

With regard to the first recommendation, it is worth noting that diagnosticians should be differentiating between specific procedures, each of which has its own particular system of interpretation, but not between individual techniques. We are therefore not asking if the TAT is valid in general terms, for there is no sensible answer to a question formulated in this manner. We may ask separately about each of the systems of interpretation, many of which have been created in the case of the TAT. They are, however, empirically grounded to varying degrees.

The second recommendation appears straightforward, but it is not so easy to comply with in Poland in the case of TAT, because there are no certified centres offering instruction in a specific manner of interpretation.

The third recommendation applies to the confirmation of instances of sexual abuse, for which diagnosticians are particularly keen to use projective drawing, and in Poland – the Tree Test. It should be noted, however, that there is a shortage of empirical data re-

garding the link between the experiencing of sexual abuse and any signs in the Tree Test and other drawing techniques (as well as in the Rorschach Test and TAT). Only the SCORS system for TAT [72] makes it possible to estimate the psychological effects of experiencing abuse.

It is worth highlighting that the authors based this recommendation on a meta-analysis of twelve studies, each of which were verified by several different projective indicators [17]. Some of the results of these studies made it possible to differentiate between children who had experienced sexual abuse and children who had not fallen victim to it, however there were also many results in which no fundamental differences in the values of the projective indicators were found between these groups. The meta-analysis provided a response to the question of whether, on taking all these results into account, it is meaningful to speak of the existence of a systematic difference between these groups. This response is indeed positive, but completely useless for practical purposes – of course, systematic confirmation is being provided of a difference, but the magnitude of this effect is negligible (the statistics make it apparent that on average, the results of one group differ from the results of the second by less than half a standard deviation). Similar results were provided by a second meta-analysis (referred to in [16]), which took a larger amount of studies into consideration. These results indicate that SCORS is the only system able to distinguish between children who are victims of sexual abuse and children who have not experienced physical abuse. All other projective indicators are of no diagnostic value in this area, so applying them for the purpose of uncovering abuse can be harmful.

The fourth recommendation urges flexibility when applying projective techniques depending on the assessment's aim and context. In certain contexts and for some research purposes, it is possible to apply standards that are not so strict. If our assessment will mainly be of benefit to ourselves or our team and if we are completing it for the purposes of psychotherapy we are conducting or, for other reasons, we are remaining in regular contact with the test subject and after some time, we will have the opportunity to ascertain the extent to which our assessment is accurate and, if need be, correct it, we can then permit the use of less reliable indicators and procedures. Nevertheless, if, as an outcome of the research, we are to issue a formal opinion for institutional purposes on which some important decisions affecting the subject in our studies will depend and if a judgement will "follow him into the world" and we will have no opportunity to correct

any potential errors, it is absolutely essential to maintain the highest standards. In such circumstances, the only systems for interpreting projective data that can be relied on are those which have been positively verified in studies and independently replicated.

This second kind of assessment is assessment for the purposes of justice. It should therefore be considered whether (and if so, which) systems of interpreting TAT are able to fulfil the high standards described above and whether, in such a case, this tool can actually be employed in forensic assessment.

3. Clinical assessment

It is possible to distinguish two main approaches to the interpretation of TAT stories – the clinical approach and the experimental approach. The primary representative of the former was the creator of TAT himself, Henry Murray³, who designed the handbook [45] and initial system for the interpretation of projective stories based on his own original theory of personality known as personology [44, see also 22, 68]. This is the most familiar system in Poland, probably because it was described in the TAT handbook and is based on recognising the needs of the protagonist and environmental pressures in the stories.

The creator of the next system was Leopold Bellack, who initially collaborated with Murray and was even listed as one of the four co-authors of the handbook for a subsequent version of the test [75]. However, after a few years, he parted company with the team and created his own approach based on ego theory. He also developed versions for children (CAT and CAT-H [6, 8]) as well as the elderly (SAT [7]). Bellack's approach was extremely popular in America in the second half of the 20th century; the last edition of this widely read book (the sixth) came out in 1997, and at the end of his life, Bellack was developing a new, revised version of the handbook [5] with the assistance of a close colleague. In Poland, Bellack's approach became known thanks to the work of Adam Frączek [14] and Anna Suchańska [67, 68].

It should be noted that both of these approaches (the two most popular) are not based on the results of academic research, but only on theoretical concepts which were not empirically verified. On the strength of these concepts, it was assumed that certain elements in projective storytelling had a particular meaning; this assumption was not however tested in systematic vali-

³ In fact, the creators of TAT were both Murray and his co-worker Christiana Morgan (Morgan and Murray, 1937).

dation studies, either by Murray or by Bellak. For the sake of clarity, it must be added that conducting tests was hardly an alien process either to Murray or Bellak. Murray [44] conducted interesting and frequently cited research on the projection phenomenon. He aroused specific emotions in the children participating in the research and went on to ask them to describe several human faces in the presented photographs. It turned out that the aroused emotions were projected, causing, for example, an increased tendency to perceive the figures in the pictures as bad or having bad intentions. By contrast, Bellak, in one of his works [3] demonstrated the influence of emotions induced by posthypnotic suggestion on the content of projected stories. These experiments demonstrated that emotions experienced in real time are reflected in the content of projected stories, thus proving the existence of the general phenomenon of projection. No proof was produced, however, of the validity of the systems of interpretation created by both authors on the basis of theory and clinical experience. In the meantime, studies conducted by other researchers conclusively demonstrated that things were very often completely different from what Bellak and Murray believed.

4. Clinical interpretations in the light of research

Bellak [4] thought that it was possible, on the basis of the TAT, to recognise suicidal tendencies. Of particular significance in this case are the stories for cards 3BM (a kneeling female figure by a bed) and 14 (a human figure in a bright window opening), because in these scenes, potential suicide victims should actually see their own acts of suicide. However, in a study purposefully devoted to this issue [31], this thesis was not confirmed. Patients with at least one suicide attempt behind them failed at interview to compose any stories about suicide at all, either for these pictures or for any others. There were also no recorded instances of the stories composed in response to these cards displaying any distinctive traits when compared to the stories of a control group and the stories composed by the same patients for two other cards. However, the study delivered an outcome unanticipated by Bellak – the suicides were distinguished by a tendency toward so-called dichotomous thinking. Specifically, their stories more frequently featured sentences containing Hamletian dilemmas of the kind “to act or not to act”, not only relating to life and death, but also to different, quite prosaic matters.

Moreover, no confirmation was obtained of Bellak’s assertion that the stories of people suffering from schizophrenia could be recognised, among other factors, from disorders of perception, stunted emotions and the appearance of bizarre content. Bellak [4, p. 139] provides “seeing without looking” as an example of the latter. However, in a study [60] purposefully devoted to this issue, the stories of people diagnosed with psychosis were compared with the stories of people who had been diagnosed with depression and those of people without a diagnosis. The first of these were not distinguished either by the unusualness of their story content, or their affective tone or even disorders of perception. Bellak’s assertions on the appearance of distinctive traits in the stories of people with ADHD [59] also turned out to be wide of the mark.

Studies have also negatively verified Bellak’s convictions on the latent meaning of individual cards. In a methodologically refined study, Nancy Alvarado [2, more detailed description in 63] established that the tableaux which Bellak regarded as representing the same motif, for example hostility between mother and daughter or the Oedipus Complex, did not appear at all similar to her study subjects.

Bellak [4] believed that academic psychologists experience unwarranted doubts with regard to projective techniques, because there is nothing irrational or mysterious in them. In the story interpretation scheme he designed, three levels of inference are taken into account: descriptive, interpretive and diagnostic. The diagnostician should progress step by step from a description of the story to complex diagnostic hypotheses.

Bellak [4, p. 450] illustrates this idea with the following example: “(...) if in response to card #1 from the TAT (a boy with a violin) a boy tells me the following story: «this boy has been forced by his mother to practice the violin. He is sitting here and thinking about breaking it» – then I formulate a conclusion close to my observation on the descriptive level that the story relates that the boy was ordered to practice the violin, but instead of this he is thinking about breaking it”. On the interpretive level: “on receiving an order from his mother to perform a task, he feels anger and readiness for aggression”. On the diagnostic level: “this person has large reserves of anger whose source can be traced to the demands made of him by women”.

However, in the progression through successive levels described above, not everything seems as obvious as Bellak thought. It is possible to imagine other potential reasons for composing such a story, for example that the young story-teller had just encountered some such model of aggressive behaviour and this

model had appealed to him. This could have been provided by an older brother or friend or film character – since the time of Bandura’s experiments, it has been understood what a strong influence such observations exert on behaviour. In such cases, Bellak’s interpretation would be an instance of the fundamental attribution error entailing the explanation of behaviours by means of internal dispositions and not external situational factors. Bellak is directly speaking in this case about feelings of anger experienced by the story-creator that are triggered by the demands of a woman. In fact, what the creator is actually experiencing may be an array of other emotions, for example joy and pride at behaviour that he imagines to be in compliance with a new, aggressive model. There could also be a whole array of other emotions.

What the aggression in the story signifies is an issue which has not been resolved in an explicit empirical manner. Both the thesis that the aggression in the story corresponds to aggression exhibited through behaviour and the opposite thesis that the aggression in the story denotes maintaining control over one’s own aggression have attracted clear support. In one study [26], for example, it turned out that more aggression can be encountered in the stories of psychology students than those of perpetrators of serious crimes employing violence.

For the sake of introducing a sense of balance into this portrayal of the clinical approach, it is worth recalling studies undertaken by Bernard Murstein [49, 50], who attempted to investigate the extent to which projected content stemmed from the research subject’s personality as opposed to the picture itself and various contextual factors (the testing circumstances, relationship with the assessor and the current status of the person being assessed et al.). It is worth highlighting that such an approach makes it possible to avoid the fundamental attribution error and perceive human behaviour in all its complexity, conditioned both by personal dispositions and situational factors as well as interaction between the two. In his studies, Murstein [48, 49] used – as example content – projected manifestations of hostility and somewhat fortuitously created an empirical diagnostic system to evaluate this variable [51]. On the basis of other studies, Murstein [46, 47] created sets of normative data⁴ relating to the appearance frequency of various kinds of themes exhibited by stories created for a given picture. As can be seen, Murstein, by investigating the theoretical bases

of his own system, created a theoretical basis for accurate interpretation, while taking care to ensure standards of some kind were met. Despite Murstein’s clinical approach being almost half a century old, it may be regarded, even today, as modern and compatible with the current trend for assessment based on empirical evidence (EBA).

5. The original TAT procedure in the light of research

Empirical studies show that the research procedure and accompanying set of instructions developed by Murray should not be applied in their original form. Why not? Because at least two elements possess disqualifying defects. The first of these is the instructions, which refer to intelligence (the test supposedly measures imagination (“which is one form of intelligence”) and achievements (“compose a story as interesting as you can manage”, “we will see how well you do”). However, it is well known that this is the kind of information that is given in psychological experiments if the tester is interested in the influence of stress or a threat to the ego on some type of human functioning. Allan Lundy’s research shows that the TAT’s validity and reliability are destroyed by any mention in the instructions of being assessed (and this is precisely how measuring of intelligence or levels of achievement are understood) and even a mediating factor in the environment, such as the assessment being carried out in the school class in which the test subjects have their lessons from day to day.

Murray was incorrect to assume that the figures in the pictures should be selected to match the characteristics of the people being assessed, thus children should be given tableaux presenting children, adults – tableaux depicting adult characters, women – tableaux presenting women, and men – tableaux featuring male protagonists. Murray [45] thought that this selection procedure would make it easier for test subjects to identify with the protagonists in their stories, which would, in Murray’s view, ensure the creation of stories that were more personal, and therefore of greater diagnostic value. However, other researchers, including Bellak [4], thought that characters different from the test subjects allowed them to distance themselves from them and project problematic content or their own unaccepted feelings, needs or desires onto them. Two researchers from David C. McClelland’s team – Edward L. Walker and John Atkinson – also took this assumption as a departure point in their research devoted to fear of a nuclear disaster in soldiers. This research was

⁴ Bellak also did this for the CAT – his handbook contains data on the frequency of various motifs in the stories of a sample of around 40 children whom he assessed.

conducted at an atomic testing ground under natural conditions and under conditions igniting fears of an explosion. The participants were asked to compose stories to go with a picture. In the first group, the figure in the picture was identified by the participants as a man, while in the second group, it was identified as a woman. It turned out that in the second group, many more fantasies about fear appeared than in the first, which lends support to the argumentation above.

Also, one study by the author of this article [63] tested for disidentification with the protagonist in the story, i.e. measures taken by the test subject targeted at making the protagonist different to his own self. It was found that characters which triggered disidentification were a “screen for the projection” of content that the participant claimed in an independent task (a Q-sort with individually ranked material) was not a close match for his actual self and even less of a match for his ideal self.

It is therefore clear that people have a great deal to say about characters that are unlike their own selves. Interesting stories are told by women about men and by men about women, as was found in one study. Frances Worchel, Lana Aaron and Deanna Yates [81] conducted an experiment in which they gave one group of women tableaux with female characters and a second group, tableaux with male characters. They acted in similar fashion with two groups of men, ultimately yielding four groups in which the main effect of the test subjects’ gender, the main effect of the “gender” of the tableau and the interaction between both of these factors could be taken into account in a variance analysis. It turned out that when women speak about men or men about women, then these stories differ from the stories in which everyone is speaking about their own sex, in that the former contain significantly more fantasies relating to anxiety and withdrawal, i.e. feelings which people are usually reluctant to recognise in themselves.

The Worchel, Aaron and Yates study is interesting for one other reason. The dependent variables in this research came from the Reuben Fine system [13]. Both main effects turned out to be significant, although each one was for different dependent variables. It is hardly a revelation (since this is familiar from other sources) that a greater interest in interpersonal relations was found in women than in men. However, the second main effect that was confirmed is extremely important for the TAT research procedure. It was discovered that the cards intended for women (i.e. with female characters) trigger more fantasies in both sexes indicative of emotional arousal, which is understood in the applied system as a manifestation of

psychopathology. It therefore turns out that the classic TAT procedure is culturally biased – the women being assessed with it are systematically being evaluated as more “disturbed”.

It is therefore evident that the TAT procedure requires thorough revision and updating (cf. [63]). It is, after all, somewhat of an aberration to employ a diagnostic tool without any alterations in the same form as it was originally created almost 80 years ago – and this is exactly what is happening with TAT, as well as with many other tests that are either projective or incorrectly classified as such [cf. 57, 65]. The articles mentioned here discuss the Tree Test and Szondi Test, among others.

6. The experimental approach

The critics’ charges and negative research results mainly relate to the clinical tradition, in particular to the conceptions of Murray and Bellak (although Murstein’s ideas stand up well in studies). However, the clinical tradition is not alone. Another powerful trend is the experimental approach. It has given birth to very broad and narrow systems of interpretation usually designed to measure very clearly defined variables. These systems are based on coherent theoretical concepts and possess good empirical grounding, so it is unsurprising that Scott Lilienfeld and his group of critically-minded colleagues speak well of them. Currently, the three most widely known systems of this kind are: David McClelland’s system designed to evaluate three needs: achievement, affiliation and power; Drew Westen’s SCORS system [70, 71], which enables selected aspects of social cognition and object relations to be estimated using scales; and the DMM Phoebe Cramer system, which enables defence mechanisms to be evaluated. These systems arose during the course of many years of research, and their validity and reliability have been confirmed.

Validity is often proven by means of planned experiments in which the introduction, through experimental manipulation, of a particular factor that accords with the theory underlying a given system should cause specific changes in the projected indicators of an assessed variable. For example, in McClelland’s team, valid projective indicators of the different needs were obtained by conducting a study involving two groups. The concept behind the study entailed arousing a given need in one (criterion) group and giving neutral tasks to the other (comparative) group; next, both groups were asked to compose stories based on the same picture and differences were sought between

these stories. This is how projective indicators were attained of the needs for achievement, affiliation and power, and these were repeatedly verified later on in subsequent studies [63, 78]. Today, it is not only known that these indicators are valid, but also that the needs they measure play a crucial role in a whole array of issues belonging to different fields of psychology. The application of the work of McClelland's team⁵ in organisational, work and business psychology [39, 41] is most widely known, but in a later period, he became interested in the influence of latent needs on different spheres of human activity and functioning. Notable publications appeared, for example, on the relation between latent needs and health (in particular, the need for power [36, 38]), the role of hidden motivation in education (especially in the art field [79]) as well as educational practices which support the development of hidden motivation (linked to achievement, affiliation and power) which – as studies have shown – is something else and develops differently from the overt motivation ascribed to oneself during a self-report⁶ [40]. Later researchers followed the research trail that McClelland staked out, using a similar empirical method to create projective indicators of many other variables, such as the need for intimacy [34] or responsibility [76, 77].

Another system of interpretation for projected stories whose validity and reliability have been thoroughly investigated is the SCORS (Social Cognition and Object Relations Scoring) system created by Drew Westen [70, 71]. Westen managed to achieve the ambitious task of integrating two “mainstream” trends – respectively, social cognition (from academic psychology) and object relations (from clinical psychology) – into one system. This was possible because both trends essentially use the same operationalisations, for example they are interested in the complexity of representations of people. This was how the first SCORS scale was named, in which a story could receive from 1 to 7 points. A lower score should be granted if the representations of people presented in a story are egocentric and it happens that the subject's own thoughts or feelings are confused with those of others. A mid-range score (3 points) is applied when the test subject describes human personalities and internal states in a minimal, simplified manner, or divides an image (most likely of the same person) into good and bad.

⁵ Besides latent needs, McClelland also gained renown as a result of his widely cited studies on the role played by personal competencies in achieving success at work [37].

⁶ This may instead be linked to latent cognition, which the creator of this notion, Anthony Greenwald, did after all mention [19].

A few more points – up to 5 – are granted when the subject's representations of himself and others are stereotypical and conventional, but he manages to integrate their good and bad sides. The highest score in this scale, i.e. 7 points, can be attained for diverse and complex representations indicating that the story-teller possesses insight into himself and others.

The other SCORS scales, eight in total, were defined in a similar fashion. They were defined as follows: complexity of representations of people, affective quality of representations, emotional investment in relationships, emotional investment in values and moral standards, the causes of social events, the experiencing and control of aggressive impulses, self-esteem, identity and cohesion of the self.

The content of these scales and scoring system were created on the basis of theory (mainly object relations theory). The first versions of the scales were slightly different in form and were modified under the influence of research data until they took the shape outlined above. The most interesting and potentially useful feature of the system under discussion is the opportunity it presents for estimating psychopathology on the basis of cognitive-affective distortions affecting the representation of social objects. At this level, the validity of SCORS has been verified, in particular its diagnostic validity. It has turned out that, to varying degrees, different scales differentiate groups with diagnosed disorders of particular types. For example, people with borderline personality disorder differed from people with other diagnoses and undiagnosed people, as they score lower on the scale measuring the affective quality of representations – their images of the characters in their stories were biased toward characters with evil intentions who intended to harm others [72]. The results were replicated in two other studies involving patients diagnosed with borderline disorder [53, 73, 74]. What is more, the possibility of using SCORS to differentiate between patients with borderline disorder and other psychiatric patients as well as healthy people has been confirmed in independent studies [1, 9].

Many studies of this kind have in fact been conducted. Some of them were devoted to validity within the context of the assessment of the effects of sexual and physical abuse suffered during childhood. In a series of studies, Sidney R. Ordooff and his colleagues [54, 55, 56] demonstrated the SCORS' potential for differentiating groups of children who had experienced sexual abuse, physical abuse or conversely, had not experienced this kind of trauma. These studies have been continued and replicated by other authors [27, 61, 69]. Additional studies focusing on reliability

and validity in various applications are still being conducted [58].

It appears that the SCORS does indeed enable the discovery and description on defined scales of distortions in social representations that usually result from the suffering of physical and sexual abuse (as well as differentiating between the two of them). This was after all actually confirmed by the critical researchers Lilienfeld, Wood and Garb [30, p. 44], although they also indicate what else needs to be investigated.

These critical authors [29] have shown less favour to the DMM (Defensive Mechanisms Manual) created by Phoebe Cramer to measure defence mechanisms. The DMM has three scales which progressively describe three increasingly mature defence mechanisms: denial, projection and identification. The system is based on a theoretical conception alluding to psychoanalysis, but goes beyond this, in particular with regard to the fact that development through life is taken into account. The system also has a strong empirical basis. So why was it not appreciated by the group of critical researchers? After thorough analysis of their argumentation [23; cf. 64], only one definite conclusion becomes apparent: Lilienfeld, Wood and Garb's negative view [30] of DMM results from the large number of errors and omissions that occurred while they were studying the literature on this system. By way of example, one of the tests of validity shall be described [11] that Lilienfeld et al. [29] erroneously regarded as a reliability test. Two groups of children told stories to pictures. In the middle of the test, they were informed that they had done well or badly in the task they had performed earlier (after the test, a meticulous debriefing was carried out). It was expected that the information about having failed would cause reinforcement of defence mechanisms, while information about being successful would not constitute a reason for this, so the difference between the two groups – but only in the second part of the test – would be evident on the DMM scales. Given the age of the test subjects, it was primarily differences in denial or projection mechanisms that were expected. The test results precisely fulfilled these expectations, which provides perfect confirmation of the DMM's validity. Cramer conducted more tests of this type, using psychological threats of various kinds to trigger defence mechanisms, for example criticism of the stories composed by the research subjects [10] or undermining gender identity by using the alleged results of other studies [11]. It must therefore be concluded that the DMM's validity with regard to measuring the level of three kinds of defence mechanism has been proven.

Lilienfeld et al. [29] do not, however, view the matter this way. In their view, the studies described above tested the DMM's stability (one of the forms of reliability) using the test-retest method. They were doubtlessly drawn to take this stance by the fact that two measurements, separated in time as it were, were being compared. However, this method is also used to investigate validity. This is the method known as the analysis of non-accidental modifications to test results [cf. 24]. The critical authors themselves also must have perceived their errors in relation to the Cramer system, because in their later publications, they failed to repeat these charges (at least, they did not refer to them – basically, they failed to express their opinion on the topic).

7. Should they be applied or not? Summary

Taking everything that has been written above into account, it can be stated that the situation regarding TAT is as follows: the most popular systems of interpretation – i.e. those of Murray and Bellak – fail to meet the requisite standards, in particular their validity is yet to be proven (in fact, studies have demonstrated the lack of validity of some popular interpretations). We should therefore not be using these systems at all, and certainly no longer in assessment for court requirements.

What, then, with these systems which have successfully survived a confrontation with empirical data – in the form of the McClelland, Westen and Cramer systems? And also with the Murstein system for assessing hostility, or many other systems, not mentioned here, for assessing specific variables?⁷

Let's take another look at the criteria included in the second part of this article. The first of these would appear to be met in relation to all the empirically grounded systems mentioned above if it were not for the requirement to possess standards. The creators of these systems failed to develop any general standards, recommending instead that local standards be created by the individual centres using a given system. In Poland, this in practice meant a need to either conduct normalisation studies for Polish populations or independently develop standards at the various diagnostic centres. Even if either of these procedures is completed, a further – possibly temporary – obstacle remains to using empirical systems for TAT interpretation in assessment for court requirements. As the second rec-

⁷ Some of these systems have been described in another work [63].

ommendation proposes, before beginning to apply a given system in one's own practice, one should undergo training in its use. However, there are currently no centres in Poland offering training in SCORS or DMM, while the McClelland system is reserved for a single firm continuing his team's traditions. Quite possibly, someone may manage, despite these constraints, to undergo training in Poland or abroad. After resolving the question of standards, this person would therefore be in a position to apply the SCORS system, for example, to the assessment of the effects of abuse, DMM to describing the intensity and maturity of defence mechanisms, or the McClelland to assessing motivation linked to needs for achievement, affiliation and power, which can be crucial, and not only within a career context (other lesser known scales also remain to be chosen).

However, the notion that TAT (or any other projective tool) is an "X-ray" enabling the psychologist to illuminate the human soul and see anything that is hidden inside must be renounced for good. Projective techniques – or any other for that matter – cannot be used to measure the complete human psyche. It should be noted that systems without empirical grounding, in particular Bellak's system, were attempting to do exactly this – to provide knowledge for the diagnostician about a whole personality, a complete pathology. Such ambitious tasks cannot be performed through scientific psychological assessment, even though laypeople, for whom the psychologist is reminiscent of a clairvoyant or fortune teller, may expect (and sometimes fear) this. Scientific methods can only be used to undertake much more modest tasks, for example describing the heightening of needs or defence mechanisms. Theory enables us to infer conclusions from this that go beyond the raw data, but not too far. And this is quite possibly the most important thing that is worth remembering with regard to projective techniques or diagnostic tools in general. If there is no way of gaining immediate access to reliable studies, and the psychologist wishes to know if a technique being presented to her is of diagnostic value, only one criterion need be considered: how much that technique promises to reveal about the test subject. If one clearly defined variable exists, or a few (or in exceptional cases, more), this technique should be given a chance, while any techniques promising insight into the deep, unconscious layers of personality and supposedly revealing diverse manifold truths about test subjects are for sorcerers and not psychologists!

References

1. Ackerman S. J., Clemence A. J., Weatherill R. [et al.], Use of the TAT in the assessment of DSM-IV Cluster B personality disorders, *Journal of Personality Assessment* 1999, 73, 422–448.
2. Alvarado N., Empirical validity of the Thematic Apperception Test, *Journal of Personality Assessment* 1994, 63, 59–79.
3. Bellak L., The concept of projection: an experimental investigation and study of the concept, *Psychiatry* 1944, 4, 353–370.
4. Bellak L., The Thematic Apperception Test, the Children's Apperception Test, and the Senior Apperception Technique in clinical use, Grune and Stratton, New York 1986.
5. Bellak L., Abrams D. M., The Thematic Apperception Test, the Children's Apperception Test, and the Senior Apperception Technique in clinical use, Allyn and Bacon, Boston 1997.
6. Bellak L., Bellak C. P., The Children's Apperception Test Human (C.A.T-H.), Larchmont, New York 1952.
7. Bellak L., Bellak C. P., Manual: Senior Apperception Test, Larchmont, New York 1973.
8. Bellak L., Bellak C. P., The Children's Apperception Test (C.A.T.), Larchmont, New York 1949.
9. Byeongseok K., The differences in cognitive complexity and affect-tone of object representations reflected in the early childhood memories between borderline personality disordered and normal subjects, *Dissertation Abstracts International. Section B: The Sciences and Engineering*, 1996, 56 (11-B), 6443.
10. Cramer P., Anger and the use of defense mechanisms in college students, *Journal of Personality* 1991, 59, 39–55.
11. Cramer P., Gaul R., The effect of success and failure on children's use of defense mechanisms, *Journal of Personality* 1988, 56, 729–742.
12. Explorations in personality: A clinical and experimental study of fifty men of college age, Murray H. A. [ed.], University Press, Oxford 1938.
13. Fine R., A scoring scheme for the TAT and other verbal projective techniques, *Journal of Projective Techniques* 1955, 19, 306–309.
14. Frączek A., Projekcyjne techniki badania osobowości, [in:] Wołoszynowa L. [ed.], Materiały do nauczania psychologii, t. 3, PWN, Warszawa 1965.
15. Garb H. N., Incremental validity and the assessment of psychopathology in adults, *Psychological Assessment* 2003, 15, 508–520.
16. Garb H. N., Wood J. M., Lillienfeld S. O. [et al.], Effective use of projective techniques in clinical practice: data help with selection and interpretation, *Professional Psychology: Research and Practice* 2002, 33, 454–463.
17. Garb H. N., Wood J. M., Nezworski M. T., Projective techniques and the detection of child sexual abuse, *Child Maltreatment*, 2000, 5, 161–168.

18. Garb H. N., Wood J. M., Nezworski M. T. [et al.], Toward a resolution of the Rorschach controversy, *Psychological Assessment* 2001, 13, 433–448.
19. Greenwald A. G., Banaji M. R., Utajone poznanie społeczne: postawy, wartościowanie siebie i stereotypy, *Przegląd Psychologiczny* 1996, 38, 11–63.
20. Groth-Marnat G., Handbook of psychological assessment, Wiley, New Jersey 2003.
21. Groth-Marnat G., Roberts L., Human Figure Drawings and House Tree Person Drawings as indicators of self-esteem. A quantitative approach, *Journal of Clinical Psychology* 1998, 54, 219–222.
22. Hall C. S., Lindsey G., Campbell J. B., Teorie osobowości, Wydawnictwo Naukowe PWN, Warszawa 2004.
23. Hibbard S., A critique of Lilienfeld et al.'s (2000) "The scientific status of projective techniques", *Journal of Personality Assessment* 2003, 80, 260–272.
24. Hornowska E., Testy psychologiczne – teoria i praktyka, Wydawnictwo Naukowe Scholar, Warszawa 2001.
25. Karon B. P., The clinical interpretation of the Thematic Apperception Test, Rorschach and other clinical data: a reexamination of statistical versus clinical prediction, *Professional Psychology: Research and Practice* 2000, 31, 230–233.
26. Keltikangas-Järvinen L., Alexithymia in violent offenders, *Journal of Personality Assessment* 1982, 46, 462–467.
27. Kernhof K., Kaufhold J., Grabhorn R., Object relations and interpersonal problems in sexually abused female patients: An empirical study with the SCORS and the IIP, *Journal of Personality Assessment* 2008, 90, 44–51.
28. Kubiszyn T. W., Meyer G. J., Finn S. E. [et al.], Empirical support for psychological assessment in clinical health care settings. *Professional Psychology: Research and Practice* 2000, 31, 119–130.
29. Lilienfeld S. O., Wood J. M., Garb H. N., The scientific status of projective techniques, *Psychological Science in the Public Interest* 2000, 1, 27–67.
30. Lilienfeld S. O., Wood J. M., Garb H. N., Status naukowy technik projekcyjnych, Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków 2002.
31. Litinsky A. M., Haslam N., Dichotomous thinking as a sign of suicide risk on the TAT, *Journal of Personality Assessment* 1998, 3, 368–378.
32. Lundy A., The reliability of the Thematic Apperception Test, *Journal of Personality Assessment* 1985, 49, 141–145.
33. Lundy A., Instructional set and Thematic Apperception Test validity, *Journal of Personality Assessment* 1988, 52, 309–320.
34. McAdams D. P., The intimacy motive, [in:] *Motivation and personality: Handbook of thematic content analysis*, Smith C. P. [ed.], Cambridge University Press, New York 1992.
35. McAdams D. P., The intimacy motivation scoring system, [in:] *Motivation and personality: Handbook of thematic content analysis*, Smith C. P. [ed.], Cambridge University Press, New York 1992.
36. McClelland D. C., Motivational factors in health and disease, *American Psychologist* 1989, 44, 675–683.
37. McClelland D. C., Testing for competence rather than for "intelligence", *American Psychologist* 1973, 28, 1–14.
38. McClelland D. C., The need for power, sympathetic activation and illness, *Motivation and Emotion* 1982, 6, 31–41.
39. McClelland D. C., Boyatzis R. E., The leadership motive pattern and long-term success in management, *Journal of Applied Psychology* 1982, 67, 737–743.
40. McClelland D. C., Koestner R., Weinberger J., How do self-attributed and implicit motives differ?, *Psychological Review* 1989, 96, 690–702.
41. McClelland D. C., Winter D. G., Motivating economic achievement, Free Press, New York 1969.
42. Meehl P. E., The cognitive activity of the clinician, *American Psychologist* 1960, 15, 19–27.
43. Motives in fantasy, action and society. A method of assessment and study, Atkinson J. W. [ed.], Van Nostrand, Princeton 1958.
44. Murray H. A., The effect of fear upon estimates of the maliciousness of other personalities, *Journal of Social Psychology* 1933, 4, 310–329.
45. Murray H. A., Test Apercepcji Tematycznej, Laboratorium Technik Diagnostycznych im. B. Zawadzkiego, Warszawa 1986.
46. Murstein B. I., A normative study of TAT ambiguity, *Journal of Projective Techniques and Personality Assessment* 1965, 28, 210–218.
47. Murstein B. I., Normative written responses for a college sample, *Journal of Projective Techniques and Personality Assessment* 1972, 36, 109–147.
48. Murstein B. I., Projection of hostility on the TAT as a function of stimulus, background and personality variables, *Journal of Consulting Psychology* 1965, 29, 43–48.
49. Murstein B. I., The effect of stimulus background, personality, and scoring system on the manifestation of hostility on the TAT, *Journal of Consulting and Clinical Psychology* 1968, 32, 355–365.
50. Murstein B. I., The stimulus, [in:] *Handbook of projective techniques*, Murstein B. [ed.], Basic Books, New York 1965.
51. Murstein B. I., David C., Fisher D. [et al.], The scaling of the TAT for hostility by a variety of scaling methods, *Journal of Consulting Psychology* 1961, 25, 497–504.
52. Murstein B. I., Mathes S., Projection on projective techniques = pathology: The problem that is not being addressed, *Journal of Personality Assessment* 1996, 66, 337–349.

53. Nigg J., Lohr N., Westen D. [et al.], Malevolent object representations in borderline personality disorder and major depression, *Journal of Abnormal Psychology* 1992, 101, 61–67.
54. Ornduff S. R., TAT assessment of object relations: Implications for child abuse, *Bulletin of the Menninger Clinic* 1997, 61, 1–15.
55. Ornduff S. R., Freedenfeld R. N., Kelsey R. M. [et al.], Object relations of sexually abused female subjects: A TAT analysis, *Journal of Personality Assessment* 1994, 63, 223–238.
56. Ornduff S. R., Kelsey R. M., Object relations of sexually and physically abused female children: A TAT analysis, *Journal of Personality Assessment* 1996, 66, 91–105.
57. Paluchowski W. J., Stemplewska-Żakowicz K., The reliability of projective techniques as tools of psychological assessment. Part 2: The Tree Test, *Problems of Forensic Sciences* 2013, 94.
58. Peters E., Hilsenroth M., Eudell-Simmons E. [et al.], Reliability and validity of the Social Cognition and Object Relations Scale (SCORS) in clinical use, *Psychotherapy Research* 2006, 16, 617–626.
59. Ponto J. M., Differential validity of the TAT in distinguishing attention deficit hyperactivity disorder from externalizing disorders, *Dissertation Abstracts International: Section B: The Sciences & Engineering* 1998, 59 (2-B), 0909.
60. Sharkey K. J., Ritzler B. A., Comparing diagnostic validity of the TAT and a new picture projection test, *Journal of Personality Assessment* 1985, 49, 406–412.
61. Slavin J. M., Stein M. B., Pinsker-Aspen J. H. [et al.], Early memories from outpatients with and without a history of childhood sexual abuse, *Journal of Loss and Trauma* 2007, 12, 435–451.
62. Stemplewska-Żakowicz K., Diagnostyka psychologiczna. Autorskie podręczniki akademickie, Strelau J., Doliński D. [eds.], Gdańskie Wydawnictwo Psychologiczne, Gdańsk 2009.
63. Stemplewska-Żakowicz K., O rzeczach widywanych na obrazkach i opowiadanych o nich historiach. TAT jako metoda badawcza i diagnostyczna, *Academica*, Warszawa 2004.
64. Stemplewska-Żakowicz K., Status naukowy krytyki testów projekcyjnych. Recenzja książki Lilienfelda, Wooda i Garba „Status naukowy technik projekcyjnych”, *Psychologia Jakości Życia* 2004, 1, 133–141.
65. Stemplewska-Żakowicz K., Paluchowski W. J., The reliability of projective techniques as tools of psychological assessment. Part 1: Why it is unjustified to describe some of them as projective?, *Problems of Forensic Sciences* 2013, 93, 421–437.
66. Suchańska A., Test Apercepcji Tematycznej H. A. Murray, [in:] *Metody projekcyjne. Tradycja i współczesność*, Sek H. [ed.], Wydawnictwo Naukowe UAM, Poznań 1984.
67. Suchańska A., Kontrowersje wokół TAT, *Przegląd Psychologiczny* 1993, 36, 239–248.
68. Suchańska A., Test Apercepcji Tematycznej. Przez analizę treści do analizy procesu, Wydawnictwo Naukowe UAM, Poznań 1994.
69. Valls G., Object relations of women sexually abused in childhood: A TAT investigation, *Dissertation Abstracts International: Section B: The Sciences & Engineering* 1999, 60 (1-B), 0378.
70. Westen D., Clinical assessment of object relations using the TAT, *Journal of Personality Assessment* 1991, 56, 56–74.
71. Westen D., Social cognition and object relation, *Psychological Bulletin* 1991, 109, 429–455.
72. Westen D., Lohr N., Silk K. R. [et al.], Object relations and social cognition in borderlines, major depressives, and normals: A Thematic Apperception Test analysis, *Psychological Assessment* 1990, 2, 355–364.
73. Westen D., Ludolph P., Lerner H. [et al.], Object relations in borderline adolescents, *Journal of American Academy of Child and Adolescent Psychiatry* 1990, 29, 338–348.
74. Westen D., Ludolph P., Silk K. [et al.], Object relations in borderline adolescents and adults: Developmental differences, *Adolescent Psychiatry* 1990, 17, 360–384.
75. White R. W., Sanford R. N., Murray H. A. [et al.], Morgan-Murray Thematic Apperception Test: Manual of directions, Harvard Psychological Clinic, Cambridge 1941.
76. Winter D. G., Responsibility, [in:] *Motivation and personality: Handbook of thematic content analysis*, Smith C. P. [ed.], Cambridge University Press, New York 1992.
77. Winter D. G., Scoring system for responsibility, [in:] *Motivation and personality: Handbook of thematic content analysis*, Smith C. P. [ed.], Cambridge University Press, New York 1992.
78. Winter D. G., The contributions of David McClelland to personality assessment, *Journal of Personality Assessment* 1998, 7, 129–145.
79. Winter D. G., McClelland D. C., Thematic analysis an empirically derived measure of the effects of liberal arts education, *Journal of Educational Psychology* 1978, 70, 8–16.
80. Wood J. M., Lilienfeld S. O., Garb H. N. [et al.], The Rorschach test in clinical diagnosis: A critical review, with a backward look at Garfield (1947), *Journal of Clinical Psychology* 2000, 56, 395–430.
81. Worchel F. T., Aaron L. L., Yates D. F., Gender bias on the Thematic Apperception Test, *Journal of Personality Assessment* 1990, 55, 593–602.

TEST APERCEPCJI TEMATYCZNEJ W ŚWIETLE BADAŃ. MOŻLIWOŚCI WYKORZYSTANIA W DIAGNOZIE DO CELÓW SĄDOWYCH*

1. Krytyka technik projekcyjnych

Znaczna część projekcyjnych technik diagnostycznych powstała w pierwszej połowie XX wieku. W kolejnych dekadach piśmiennictwo na temat najbardziej znanych spośród nich, takich jak test Rorschacha czy Test Apercepcji Tematycznej (TAT), szło w dziesiątki pozycji rocznie. Jednakże w latach 90. XX wieku i na przełomie wieków coś zaczęło się zmieniać. Techniki projekcyjne stały się przedmiotem dyskusji i gorących sporów, a nowe badania doprowadziły do zakwestionowania ich trafności i użyteczności praktycznej [16, 18, 23, 25, 28, 30, 80, por. 64]. Zapewne z powodu tej krytyki opartej na wnioskach z nowych badań zainteresowanie technikami projekcyjnymi zdaje się obecnie maleć. Wyszukiwanie w bazach EBSCO publikacji, które mają nazwę „Thematic Apperception Test” w tytule, przyniosło 628 pozycji opublikowanych w naukowych pismach recenzowanych w latach 1935–2012, jednak ograniczenie wyszukiwania do ostatnich 10 lat dało tylko 33 publikacje (większość w jednym czasopiśmie – *Journal of Personality Assessment*). Innym znakiem nowych czasów jest fakt, że w kolejnej edycji swego bardzo popularnego podręcznika diagnozy klinicznej Gary Groth-Marnat [20] zrezygnował z rozdziału na temat rysunku projekcyjnego z powodu narastających doniesień empirycznych o braku trafności opartych na nim wskaźników¹.

Główne zarzuty krytyków [30, 80] pod adresem technik projekcyjnych to właśnie brak trafności i rzetelności², a zamiast tego posługiwanie się impresyjnymi interpretacjami, które nie są odnoszone do żadnych norm, ponieważ z reguły norm także brak. Od dawna też podnoszony jest zarzut dopatrywania się psychopatologii w materiale projekcyjnym, przy czym im większa w sensie ilościowym jest projekcyjna „produkcja” danej osoby, tym więcej zaburzeń zostaje u niej „rozpoznanych”. Przed takim błędem we wnioskowaniu szczególnie przestrzegał Bernard Murstein [52], którego zasadniczo można zaliczyć nie do krytyków, lecz zwolenników technik projekcyjnych. W konsekwencji wymienionych słabości techni-

kom projekcyjnym brakuje trafności przyrostowej (ang. incremental validity), czyli nie obserwuje się poprawy trafności całej diagnozy po udostępnieniu diagnoście dodatkowo danych z techniki projekcyjnej [15]. Skoro wiadomo, że techniki projekcyjne nie mogą – wobec wymienionych mankamentów – prowadzić do trafnych wniosków, zatem ich stosowanie w diagnozie psychologicznej należy, zdaniem krytyków, uznać za etycznie wątpliwe. Co zatem zrobić?

2. Rekomendacje praktyczne do stosowania technik projekcyjnych

Jeśli chodzi o skłonność do przeceniania psychopatologii, bardzo możliwe, iż jest ona „w oku patrzącego” i należałoby ją uznać za jeden z błędów nieformalnego wnioskowania diagnostycznego, przed którymi przestrzegał Meehl [42]. Radą na to jest odejście od polegania jedynie na doświadczeniu czy intuicji klinicznej diagnosty i oparcie diagnostycznych sądów na stwierdzeniu – bądź nie – obecności „twardych” wskaźników psychopatologii. Tylko jakie wskaźniki możemy uznać za „twarde”? Odpowiadają na to Garb, Wood, Lilienfeld i Nezworski [16], formułując trzy warunki, które powinien spełniać wskaźnik, aby na jego podstawie można było odpowiedzialnie orzekać o psychopatologii. Warunki te są następujące:

1. wskaźnik musi mieć widoczny związek z danym zaburzeniem, cechą lub syndromem;
2. dowodzące tego dane powinny być uzyskane w badaniach poprawnych metodologicznie;
3. wyniki powinny być niezależnie zreplicowane.

Zatem nie wystarczy, że jakiś związek został wykazany w jednym badaniu. Jak podają autorzy [16], uzyskano już bardzo dużo pozytywnych wyników dotyczących tego czy innego wskaźnika projekcyjnego. Nie powinno to dziwić, zważywszy na wielką liczbę wykonanych do tej pory badań nad technikami projekcyjnymi. Jednak warte uwzględnienia w praktyce są tylko te wskaźniki, które zostały niezależnie potwierdzone w badaniach innego zespołu i ich trafność nie budzi wątpliwości. Trzeba w tym miejscu powiedzieć, że w Polsce stosuje się niestety bardzo dużo wskaźników projekcyjnych (w różnych technikach), które takiego empirycznego ugruntowania nie mają. Wskaźniki te powinny być bezzwłocznie wycofane. Niektóre z nich (dotyczące opowiadań projekcyjnych Teście Apercepcji Tematycznej) zostaną wskazane w dalszych częściach artykułu.

* Referat został wygłoszony podczas konferencji „Warunki wykorzystywania metod projekcyjnych w psychologicznej diagnozie dla potrzeb sądu”, która odbyła się w dniach 16–17 listopada 2012 roku w Lublinie. Wykorzystano w nim niewielkie fragmenty innych prac autorki [62, 63, 64].

¹ Sam był autorem jednego z takich badań [21].

² A w niektórych przypadkach nawet brak badań na ten temat, dowodzący beztróski autorów techniki co do podstawowych wymagań stawianych narzędziom diagnostycznym [por. 57, 65].

W różnych pracach krytycznych autorów zawarte są także wskazówki i zalecenia dla diagnostów stosujących techniki projekcyjne. Autorka zebrała je razem i przedstawiła [62, 63] w postaci czterech rekomendacji, które są następujące:

1. używaj procedur (systemów interpretacji), których trafność i rzetelność zostały wystarczająco potwierdzone w badaniach i dla których zostały opracowane normy;
2. przejdź trening w stosowaniu danej procedury, zanim zaczniesz ją stosować w swojej praktyce;
3. nie stosuj technik projekcyjnych do wykrywania nadużyć seksualnych lub fizycznych;
4. stosuj techniki projekcyjne w różny sposób, w zależności od celu badania.

W rekomendacji pierwszej warto zauważyć, że diagności rozróżniać mają między konkretnymi procedurami, z których każda ma swój specyficzny system interpretacji, nie zaś pomiędzy poszczególnymi technikami. Nie pytamy więc, czy Test Apercepcji Tematycznej jest trafny ogólnie, bo nie ma sensownej odpowiedzi na tak sformułowane pytanie. Pytać możemy osobno o każdy z systemów interpretacji, których w przypadku Testu Apercepcji Tematycznej (TAT) powstało już całkiem sporo. Są one jednak w niejednakowym stopniu ugruntowane empirycznie.

Rekomendacja druga wydaje się oczywista, choć nie tak łatwa do spełnienia w Polsce w przypadku TAT, ponieważ brakuje certyfikowanych ośrodków kształcących w konkretnym sposobie interpretacji.

Rekomendacja trzecia dotyczy potwierdzania nadużyć seksualnych, do czego szczególnie chętnie używany jest rysunek projekcyjny, a w Polsce Test Drzewa. Trzeba jednak stwierdzić, że brak jest danych empirycznych dotyczących związku pomiędzy doznaniem nadużyć seksualnych a jakimikolwiek oznakami w Teście Drzewa i w innych technikach rysunkowych (również w teście Rorschacha i TAT). Jedynie w systemie SCORS dla TAT [72] można oszacować psychiczne skutki doznania nadużyć.

Warto podkreślić, że rekomendację tę autorzy oparli na metaanalizie dwunastu badań, z których każde weryfikowało kilka różnych wskaźników projekcyjnych [17]. Część wyników tych badań pozwalała różnicować pomiędzy dziećmi, które doświadczyły nadużyć seksualnych i dziećmi, które nie padły ich ofiarą, jednak wiele było też wyników, w których nie stwierdzono istotnych różnic pomiędzy tymi grupami w wartościach wskaźników projekcyjnych. Metaanaliza przyniosła odpowiedź na pytanie, czy biorąc pod uwagę wszystkie te wyniki, można mówić o istnieniu systematycznej różnicy między owymi grupami. Odpowiedź jest wprawdzie pozytywna, ale zupełnie bezużyteczna do celów praktycznych – owszem, systematycznie stwierdza się różnicę, jednak wielkość tego efektu jest pomijalna (statystyki

informują, że wyniki jednej grupy średnio odbiegają od wyników tej drugiej o mniej niż pół odchylenia standardowego). Podobne wyniki przyniosła druga metaanaliza (wzmiankowana w [16]) uwzględniająca większą liczbę badań. Jej wyniki wskazały na system SCORS jako jedyny, który potrafi rozróżniać między dziećmi będącymi ofiarami nadużyć seksualnych a dziećmi, które nadużyć nie przeżyły. Wszystkie inne wskaźniki projekcyjne nie posiadały wartości diagnostycznej w tym obszarze, więc stosowanie ich w celu wykrywania nadużyć może być szkodliwe.

Rekomendacja czwarta zachęca do elastyczności w stosowaniu technik projekcyjnych w zależności od celu i kontekstu diagnozy. W pewnych kontekstach i przy niektórych celach badań można zastosować nie tak ściśle standardy. Jeśli nasza diagnoza będzie służyła przede wszystkim nam samym lub naszemu zespołowi, jeśli dokonujemy jej na użytek prowadzonej przez nas psychoterapii lub też z innych powodów pozostajemy w stałym kontakcie uczestnikiem badania i po pewnym czasie będziemy mieli możliwość przekonać się, na ile nasza diagnoza była trafna oraz ewentualnie ją skorygować, wówczas możemy dopuścić mniej pewne wskaźniki i procedury. Jednakże jeśli w wyniku badań mamy wydać formalne opinie do celów instytucjonalnych, od którego będą zależały jakieś ważne decyzje dotyczące losów uczestnika naszych badań, jeśli orzeczenie to „pójdzie w świat” za nim, a my nie będziemy mieli możliwości skorygowania ewentualnych pomyłek, to trzeba koniecznie zachować najwyższe standardy. W takich sytuacjach trzeba polegać jedynie na takich systemach interpretacji danych projekcyjnych, które zostały pozytywnie zweryfikowane w badaniach oraz przeprowadzono niezależne replikacje.

Tym drugim rodzajem diagnozy jest diagnoza na potrzeby wymiaru sprawiedliwości. Należy się zatem zastanowić, czy i ewentualnie które systemy interpretacji TAT mogą spełnić opisane wysokie wymagania i czy w takim razie narzędzie to może w ogóle być wykorzystywane w diagnozie sądowej.

3. Podejście kliniczne

Można wyróżnić dwa główne podejścia do interpretacji opowiadań TAT – podejście kliniczne i eksperymentalne. Przedstawicielem tego pierwszego był przede wszystkim sam twórca TAT Henry Murray³, który opracował podręcznik [45] i wyjściowy system interpretacji opowiadań projekcyjnych oparty na autorskiej teorii osobowości znanej jako personologia [44, zob. też 22, 68]. System ten jest w Polsce najbardziej znany chyba dlatego, że został opisany w podręczniku do TAT i opiera się na rozpoznaniu w opowiadaniach potrzeb bohatera i nacisków środowiska.

³ W zasadzie autorami TAT byli na równi Murray i jego współpracowniczka Christiana Morgan (Morgan i Murray, 1937).

Twórcą kolejnego systemu jest Leopold Bellak, który początkowo współpracował z Murrayem i nawet został wymieniony jako jeden z czterech współautorów podręcznika do kolejnej wersji testu [75]. Jednak po kilku latach pożegnał się z zespołem i stworzył własne podejście oparte na teorii ego. Opracował też wersje dla dzieci (CAT i CAT-H [6, 8]) oraz dla osób starszych (SAT [7]). Podejście Bellaka było niezwykle popularne w Ameryce w drugiej połowie XX wieku; ostatnie, szóste wydanie jego poczytnej książki ukazało się w roku 1997, przy czym pod koniec życia Bellak korzystał z pomocy bliskiego współpracownika w opracowywaniu nowych, zrewidowanych wersji podręcznika [5]. W Polsce podejście Bellaka stało się znane dzięki pracom Adama Frączka [14] i Anny Suchańskiej [67, 68].

Trzeba stwierdzić, że oba te najbardziej popularne podejścia nie opierają się na wynikach badań naukowych, lecz jedynie na koncepcjach teoretycznych, które nie były empirycznie weryfikowane. Na mocy tych koncepcji założono, że pewne elementy w opowiadaniu projekcyjnym mają określone znaczenie; nie było to jednak sprawdzane w systematycznych badaniach walidacyjnych ani przez Murraya, ani przez Bellaka. Gwoli ścisłości trzeba dodać, że ani Murrayowi, ani Bellakowi eksperymentowanie nie było całkiem obce. Murray [44] przeprowadził interesujące i często cytowane badanie nad zjawiskiem projekcji. Wzbudzał określone emocje u uczestniczących w badaniu dzieci i następnie prosił o opisanie kilku twarzy ludzkich na przedstawionych fotografiach. Okazało się, że wzbudzone emocje ulegały projekcji, na przykład strach powodował wzrost tendencji do spostrzegania postaci na obrazkach jako złych, mających złe intencje. Z kolei Bellak w jednej ze swych prac [3] zademonstrował wpływ emocji indukowanych przez posthipnotyczną sugestię na treść opowiadań projekcyjnych. Eksperymenty te pokazywały, że przeżywane aktualnie emocje odzwierciedlają się w treści opowiadań projekcyjnych, a więc dowodziły ogólnego zjawiska projekcji. Nie dowodziły jednak trafności systemów interpretacji stworzonych przez obu autorów na podstawie teorii i doświadczenia klinicznego. Tymczasem badania przeprowadzone przez innych badaczy pokazały dowodnie, że bardzo często jest całkiem inaczej niż sądzili Bellak i Murray.

4. Interpretacje kliniczne w świetle badań

Bellak [4] uważał, że na podstawie TAT można rozpoznać skłonności samobójcze. Szczególne znaczenie mają tu mieć opowiadania do tablicy 3BM (klęcząca przy łóżku postać) i 14 (sylwetka ludzka w jasnym otworze okiennym), ponieważ w scenach tych potencjalni samobójcy powinni widzieć właśnie akt samobójstwa. Jednak w specjalnie temu problemowi poświęconym badaniu [31] teza ta nie potwierdziła się. Pacjenci z co najmniej

jedną próbą samobójczą w wywiadzie wcale nie układali opowiadań o samobójstwie ani do tych obrazków, ani do żadnych innych. Nie odnotowano też żadnych szczególnych cech opowiadań w odpowiedzi na te karty w porównaniu z opowiadaniem grupy kontrolnej i z opowiadaniem układanymi przez tychże pacjentów do dwóch innych kart. Badanie przyniosło natomiast rezultat nieprzewidywany przez Bellaka – samobójcy wyróżniali się tendencją do tzw. dychotomicznego myślenia. Konkretnie w ich opowiadaniach częściej występowały zdania zawierające hamletowskie dylematy w rodzaju: „zrobić to albo nie zrobić” dotyczące nie tylko życia i śmierci, ale też różnych, całkiem prozaicznych spraw.

Nie potwierdziło się też twierdzenie Bellaka, iż opowiadania osób cierpiących na schizofrenię dają się rozpoznać m.in. po zaburzeniach percepcji, osłabieniu afektu i występowaniu dziwacznych treści. Jako przykład tych ostatnich Bellak [4, s. 139] podaje „widzenie bez patrzenia”. Jednakże w specjalnie poświęconym temu zagadnieniu badaniu [60] opowiadania osób z diagnozą psychozy porównano z opowiadaniem osób, u których zdiagnozowano depresję i z opowiadaniem osób bez diagnozy. Te pierwsze nie wyróżniały się ani niezwykłością treści opowiadań, ani ich tonem afektywnym, ani też zaburzeniami percepcji. Nietrafne okazały się też twierdzenia Bellaka na temat charakterystycznych cech opowiadań osób z zaburzeniem ADHD [59].

Negatywnie zweryfikowano też w badaniach przekonania Bellaka na temat znaczenia latentnego poszczególnych kart. W metodologicznie wyrafinowanym badaniu Nancy Alvarado [2, bliższy opis w 63] ustaliła, że tablice, które Bellak uważał za reprezentujące ten sam motyw, np. wrogość między matką a córką czy kompleks Edypa, uczestnikom badań wcale nie wydają się podobne.

Bellak [4] uważał, że psycholodzy akademicy niepotrzebnie mają wątpliwości co do technik projekcyjnych, bo nie ma w nich nic irracjonalnego czy tajemniczego. W zaprojektowanym przez niego arkuszu do interpretacji opowiadań uwzględnia się trzy poziomy wnioski: opisowy, interpretacyjny i diagnostyczny. Diagnosta powinien przechodzić krok po kroku od opisu opowiadania do złożonych hipotez diagnostycznych.

Bellak [4, s. 450] ilustruje tę myśl następującym przykładem: „(...) jeśli w odpowiedzi na kartę #1 z TAT (chłopiec ze skrzypcami) chłopiec opowiada mi następującą historię: «temu chłopcu matka kazała ćwiczyć na skrzypcach. Siedzi tu i myśli o tym, żeby je połamać» – to sformułuję wniosek bliski obserwacji na poziomie opisowym, że historia opowiada o tym, jak chłopcu kazano ćwiczyć na skrzypcach, a on zamiast tego myśli o połamaniu ich”. Na poziomie interpretacyjnym: „otrzymując od matki polecenie wykonania zadania, czuje złość i gotowość do przemocy”. Na poziomie diagnostycznym: „ta osoba ma duże pokłady gniewu, którego źródłem są stawiane jej wymagania kobiet”.

Jednak w opisanym wyżej przejściu kolejnych poziomów nie wszystko wydaje się tak oczywiste, jak sądził Bellak. Można sobie wyobrazić inne potencjalne powody do ułożenia takiego opowiadania, na przykład takie, że młody autor właśnie zapoznał się z jakimś modelem takiego agresywnego zachowania i model ten spodobał mu się. Mógł to być starszy brat czy kolega albo bohater filmowy – od czasów eksperymentów Bandury wiadomo, jak silny wpływ wywierają takie obserwacje na zachowanie. W takich przypadkach interpretacja Bellaka byłaby przykładem podstawowego błędu atrybucji polegającego na wyjaśnianiu zachowań wewnętrznymi dyspozycjami a nie zewnętrznymi czynnikami sytuacyjnymi. Bellak mówi tu wręcz o przeżywanych przez autora uczuciach gniewu wywołanego przez wymaganie kobiety. Tymczasem tym, co autor w rzeczywistości przeżywa, może być szereg innych emocji, np. radość i duma z wyobraźniowego zachowania zgodnego z nowym, agresywnym wzorem. Może to być też cały szereg innych emocji.

Co oznacza agresja w opowiadaniu, to problem, który nie doczekał się jednoznacznego rozstrzygnięcia empirycznego. Zarówno teza, iż agresja w opowiadaniu odpowiada agresji w zachowaniu, jak i przeciwna, że agresja w opowiadaniu oznacza panowanie nad własną agresją w życiu, zyskały pewne wsparcie. W jednym z badań na przykład [26] okazało się, że więcej agresji można odnaleźć w opowiadaniach studentów psychologii niż sprawców ciężkich przestępstw z użyciem przemocy.

Dla zrównoważenia tego obrazu podejścia klinicznego warto przypomnieć badania Bernarda Mursteina [49, 50], który starał się dociec, w jakim stopniu treści projekcyjne wypływają z osobowości uczestnika badania, a w jakim z samego obrazka oraz z różnych czynników kontekstowych (sytuacja testowania, relacja z badającym, aktualny stan osoby badanej itp.). Godne podkreślenia jest, że takie podejście pozwala uniknąć podstawowego błędu atrybucji i spostrzegać zachowanie człowieka w całej jego złożoności jako uwarunkowane zarówno osobistymi dyspozycjami, jak i czynnikami sytuacyjnymi oraz interakcją tych dwu. W swych badaniach Murstein [48, 49] posługiwał się – jako przykładową treścią – projekcyjnymi przejawami wrogości i niejako przy okazji stworzył empiryczny system diagnostyczny do oceny tej zmiennej [51]. Na podstawie innych badań Murstein [46, 47] stworzył zestawy danych normatywnych⁴ dotyczących częstości pojawiania się różnego rodzaju tematów opowiadań na temat danego obrazka. Jak widać Murstein, badając podstawy teoretyczne swego systemu, stworzył empiryczne podstawy trafnej interpretacji i zadbał o pewnego rodzaju normy. Pomimo że podejście kliniczne Mursteina liczy pół wieku, to można nawet dziś uznać

za nowoczesne i wpisujące się w nurt diagnozy opartej na dowodach empirycznych (EBA).

5. Oryginalna procedura TAT w świetle badań

Badania empiryczne dowodzą, że opracowana przez Murraya procedura badania wraz z instrukcją nie powinny być stosowane w oryginalnej postaci. Dlaczego? Ponieważ co najmniej dwa elementy mają dyskwalifikujące wady. Pierwszy to instrukcja, w której są wzmianki o inteligencji (test jakoby mierzy wyobraźnię „będącą jedną z form inteligencji”) i osiągnięciach („ułoż tak interesujące opowiadanie, jak tylko potrafisz”, „zobaczmy jak ci pójdzie”). Wiadomo zaś, że właśnie tego rodzaju informacje podaje się w eksperymentach psychologicznych, jeśli badaczy interesuje wpływ stresu czy zagrożenia dla ego na jakiś rodzaj funkcjonowania człowieka. Z badań Allana Lundy’ego [32, 33] wynika, że trafność i rzetelność TAT rujnuje jakakolwiek wzmianka w instrukcji o byciu ocenianym (a pomiar inteligencji czy poziomu osiągnięć tak właśnie jest odbierany), a nawet pośrednia przesłanka w otoczeniu, taka jak przeprowadzanie badań w klasie szkolnej, w której uczestnicy badania na co dzień mają zajęcia.

Nie miał racji Murray, sądząc, że trzeba dobrać postaci na obrazkach do właściwości osób badanych, a więc dzieciom dawać tablice przedstawiające dzieci, dorosłym – tablice z postaciami dorosłych, kobietom – tablice przedstawiające kobiety, a mężczyznom – tablice z postaciami mężczyzn. Murray [45] uważał, że dzięki takiemu doborowi łatwiej będzie uczestnikom badania identyfikować się z bohaterami ich opowiadań, co z kolei miało w przekonaniu Murraya zapewnić tworzenie opowiadań bardziej osobistych, a więc i bardziej diagnostycznych. Inni badacze jednak, w tym Bellak [4] sądzili, że postać niepodobna do uczestnika badania pozwala mu zdystansować się do niej i rzutować na nią treści problematyczne czy też własne nieakceptowane uczucia, potrzeby i dążenia. Z takiego założenia wyszli też dwaj badacze z zespołu Davida C. McClellanda – Edward L. Walker i John Atkinson w swoim badaniu poświęconym strachowi przed katastrofą atomową u żołnierzy. Badanie to zostało przeprowadzone na poligonie atomowym w warunkach naturalnych i wzbudzających strach przed wybuchem. Uczestników poproszono o ułożenie opowiadania do obrazka, przy czym w jednej grupie postać na obrazku była identyfikowana przez uczestników jako mężczyzna, a w drugiej grupie jako kobieta. Okazało się, że w drugiej grupie pojawiło się znacznie więcej fantazji na temat strachu niż w pierwszej, co jest wsparciem dla powyższego rozumowania.

Również w jednym z badań autorki niniejszego artykułu [63] sprawdzano dezidentyfikację z bohaterem opowiadania, czyli zabiegi uczestnika badania, mające na celu uczynienie bohatera niepodobnym do własnego Ja.

⁴ Zrobił to także Bellak dla CAT – jego podręcznik zawiera dane na temat frekwencji różnych wątków opowiadań w zbieranej przez niego próbie około 40 dzieci.

Stwierdzono, że postaci, z którymi wystąpiła dezidentyfikacja, są „ekranem do projekcji” treści, co do których uczestnik stwierdza w niezależnym zadaniu (Q-sort z indywidualnie dobieranym materiałem), że w niewielkim stopniu pasują do jego Ja realnego, a jeszcze mniej do Ja idealnego.

Widać więc, że o postaciach niepodobnych do własnego Ja ludzie mają wiele do powiedzenia. Interesujące historie opowiadają kobiety o mężczyznach, a mężczyźni o kobietach, jak to się okazało w pewnym badaniu. Frances Worchel, Lana Aaron i Deanna Yates [81] przeprowadziły eksperyment, w którym jednej grupie kobiet dały tablice z postaciami kobiecymi, a drugiej z postaciami męskimi. Podobnie postąpiły z dwiema grupami mężczyzn, otrzymując ostatecznie cztery grupy, w których w analizie wariancji można było uwzględnić efekt głównej płci uczestników badania, efekt główny „płci” tablicy oraz interakcję obu wymienionych. Okazało się, że gdy kobiety opowiadają o mężczyznach lub też mężczyźni o kobietach, to opowiadania te różnią się od opowiadań, w których wszyscy opowiadają o własnej płci tym, że zawierają istotnie więcej fantazji na temat lęku i wycofania, a więc uczuć, które zwykle ludzie niechętnie rozpoznają w samych sobie.

Badanie Worchel, Aaron i Yates jest interesujące z jeszcze jednego powodu. Zmienne zależne pochodziły w tym badaniu z systemu Reubena Fine’a [13]. Oba efekty główne okazały się istotne, choć każdy dla innych zmiennych zależnych. Nie jest rewelacją, że potwierdzono znane skądinąd większe zainteresowanie relacjami interpersonalnymi u kobiet niż u mężczyzn. Natomiast drugi stwierdzony efekt główny jest niezwykle ważny dla procedury badania TAT. Otóż wykryto, że tablice przeznaczone dla kobiet (z postaciami kobiecymi) wywołują u obu płci więcej fantazji świadczących o pobudzeniu emocjonalnym, co w zastosowanym systemie jest rozumiane jako przejaw psychopatologii. Tak więc okazuje się, że klasyczna procedura TAT jest stronnicza kulturowo – badane zgodnie z nią kobiety są systematycznie oceniane jako bardziej „zaburzone”.

Widoczne jest więc, że procedura TAT wymaga głębokiej rewizji i uwspółcześnienia (por. [63]). Jest to zresztą swoisty ewenement, by stosować narzędzie diagnostyczne bez żadnych zmian w takiej formie, w jakiej zostało ono stworzone prawie 80 lat temu – a tak właśnie rzecz się ma z TAT, a także z wieloma innymi testami projekcyjnymi lub niesłusznie do projekcyjnych zaliczanymi [por. 57, 65]. W przytoczonych tu artykułach mowa jest m.in. o Teście Drzewa i Teście Szondiego.

6. Podejście eksperymentalne

Zarzuty krytyków i negatywne wyniki badań dotyczą głównie tradycji klinicznej, w tym zwłaszcza koncepcji Murraya i Bellaka (natomiast idee Mursteina bronią się

w badaniach). Jednak tradycja kliniczna nie jest jedyna, innym potężnym nurtem jest podejście eksperymentalne. W jego ramach powstało bardzo dużo szerszych i węższych systemów interpretacji przeznaczonych zwykle do pomiaru bardzo konkretnie zdefiniowanych zmiennych. Systemy te opierają się na spójnych koncepcjach teoretycznych i mają dobre ugruntowanie empiryczne, nic więc dziwnego, że i Scott Lilienfeld wraz z grupą swych krytycznych współpracowników wyraża się o nich dobrze. Najszerzej znane obecnie są trzy takie systemy: system Davida McClellanda przeznaczony do oceny trzech potrzeb: osiągnięć, afiliacji i władzy; system SCORS Drew Westena [70, 71] pozwalający na skalach szacować wybrane aspekty poznania społecznego i relacji z obiektem oraz system DMM Phoebe Cramer pozwalający diagnozować mechanizmy obronne. Systemy te powstały w toku długoletnich badań, a ich trafność i rzetelność zostały potwierdzone.

Trafności często dowodzono poprzez zaplanowane eksperymenty, w których – zgodnie z teorią leżącą u podłoża danego systemu – wprowadzenie w drodze manipulacji eksperymentalnej określonego czynnika powinno spowodować określone zmiany w projekcyjnych wskaźnikach diagnozowanej zmiennej. Na przykład w zespole McClellanda, aby otrzymać trafne projekcyjne wskaźniki różnych potrzeb, przeprowadzano badanie z udziałem dwu grup. Idea polegała na tym, że w jednej grupie (kryterialnej) wzbudzano daną potrzebę, a w drugiej (porównawczej) dawano zadanie neutralne; następnie obie grupy przoszono o ułożenie opowiadania na temat tego samego obrazka i poszukiwano różnic między tymi opowiadaniem. Tak uzyskano projekcyjne wskaźniki potrzeby osiągnięć, afiliacji i władzy, wielokrotnie później weryfikowane w kolejnych badaniach [63, 78]. Dziś wiadomo nie tylko, że wskaźniki te są trafne, ale też że mierzone przez nie potrzeby odgrywają istotną rolę w całym szeregu problemów należących do różnych dziedzin psychologii. Najlepiej znane jest zastosowanie prac zespołu McClellanda⁵ w psychologii organizacji, pracy i przedsiębiorczości [39, 41], jednak w późniejszym okresie zainteresował się on także wpływem utajonych potrzeb na inne obszary działalności i funkcjonowania człowieka. Znaczące publikacje powstały m.in. na temat związku między utajonymi potrzebami a zdrowiem (zwłaszcza potrzebą władzy [36, 38]), na temat roli utajonej motywacji w edukacji (zwłaszcza w dziedzinie sztuki [79]), a także na temat praktyk wychowawczych, które wspierają rozwój utajonej motywacji (osiągnięć, afiliacji i władzy), która – jak pokazały badania – jest czymś innym i rozwija się inaczej niż motywacja jawna przypisywana

⁵ Oprócz utajonych potrzeb McClelland zasłynął też z szeroko cytowanych badań nad rolą kompetencji osobistych w osiągnięciu sukcesu w pracy [37].

sobie w samoopisie⁶ [40]. Kolejni badacze, podążając ścieżką wytyczoną przez McClellanda, stworzyli w podobny, empiryczny sposób projekcyjne wskaźniki wielu innych zmiennych, takich jak potrzeba intymności [34] czy odpowiedzialność [76, 77].

Kolejnym systemem interpretacji opowiadań projekcyjnych, którego trafność i rzetelność dokładnie zbadano, jest system SCORS (Social Cognition and Object Relations Scoring) stworzony przez Drew Westena [70, 71]. Westenowi udało się ambitne zadanie zintegrowania w jednym systemie dwóch „mainstreamowych” nurtów (odpowiednio w psychologii akademickiej – poznanie społeczne – i w psychologii klinicznej – relacje z obiektem). Było to możliwe, ponieważ w gruncie rzeczy nury te posługują się podobnymi operacjonalizacjami, na przykład interesuje je złożoność reprezentacji ludzi. Tak właśnie nazwano pierwszą skalę SCORS, w której opowiadanie może uzyskać od 1 do 7 punktów. Dolna punktacja powinna być przyznana, jeśli reprezentacje ludzi przedstawione w opowiadaniu są egocentryczne i zdarza się, że myśli lub uczucia własne i innych są ze sobą mylone. Punktacja pośrednia (3 punkty) jest wykorzystywana, gdy uczestnik badania opisuje ludzkie osobowości i stany wewnętrznych w sposób minimalny, upraszczający, albo rozdziela obrazy tej samej (najprawdopodobniej) osoby na dobre i złe. Nieco więcej – aż 5 punktów – jest przyznawanych, gdy reprezentacje siebie i innych są stereotypowe, konwencjonalne, ale podmiot potrafi integrować ich dobre i złe aspekty. Najwyższą punktację w tej skali, czyli 7 punktów, można uzyskać za reprezentacje różnicowane i złożone, wskazujące na posiadanie przez ich autora wglądu w siebie i innych.

W podobny sposób opisane są pozostałe skale SCORS, których w sumie jest osiem. Zostały one określone w następujący sposób: złożoność reprezentacji ludzi, afektywna jakość reprezentacji, inwestowanie emocjonalne w związku, inwestowanie emocjonalne w wartości i standardy moralne, rozumienie przyczyn zdarzeń społecznych, doświadczanie i kontrola impulsów agresywnych, samoocena, tożsamość i spójność ja.

Zawartość treściowa tych skal, a także punktacje, zostały stworzone na podstawie teorii (głównie teorii relacji z obiektem). Pierwsze wersje skal miały nieco inną postać, pod wpływem danych z badań modyfikowano je, aż przyjęły powyższy kształt. Tym, co najbardziej interesujące i potencjalnie użyteczne omawianym systemie, jest możliwość szacowania psychopatologii na podstawie zniekształceń poznawczo-afektywnych dotyczących reprezentacji obiektów społecznych. Na tym poziomie weryfikowano trafność SCORS, głównie trafność diagnostyczną. Okazało się, że różne skale w niejednakowym stopniu różnicują grupy ze zdiagnozowanymi zaburze-

niami określonego typu. Na przykład osoby z zaburzeniem osobowości typu *borderline* odróżniały się od osób z innymi rozpoznaniem i od osób bez diagnozy niższymi wynikami na skali afektywnej jakości reprezentacji – ich wizerunki postaci z opowiadań były przesunięte w stronę postaci o złych intencjach, mających zamiar szkodzić innym [72]. Wyniki te powtórzyły się w kolejnych dwóch badaniach z udziałem pacjentów z diagnozą zaburzenia pogranicznego [53, 73, 74]. Co więcej, możliwość różnicowania za pomocą SCORS pomiędzy pacjentami z zaburzeniem *borderline* i innymi pacjentami psychiatrycznymi oraz osobami zdrowymi została potwierdzona w niezależnych badaniach [1, 9].

Przeprowadzono ich naprawę sporo. Część z nich była poświęcona trafności w kontekście oceny skutków nadużyć seksualnych i fizycznych doznanych w dzieciństwie. W serii badań Sidney R. Ornduff ze współpracownikami [54, 55, 56] zademonstrował możliwości SCORS w różnicowaniu grup dzieci, które doświadczyły nadużyć seksualnych, nadużyć fizycznych, bądź też nie doświadczyły tego rodzaju traum. Badania te były kontynuowane i replikowane przez innych autorów [27, 61, 69]. Wciąż są też prowadzone dodatkowe badania na temat rzetelności i trafności w różnych zastosowaniach [58].

Wydaje się, że SCORS rzeczywiście umożliwia wykrywanie i opisywanie na skalach specyficznych zniekształceń w reprezentacjach społecznych, które zazwyczaj są skutkami doznanej przemocy fizycznej i seksualnej (różnicuje też między tymi dwiema). Przyznali to zresztą samo krytyczni badacze Lilienfeld, Wood i Garb [30, s. 44], choć zarazem wskazują, co jeszcze pozostaje do zbadania.

Mniej łaskawie krytyczni autorzy [29] obeszlili się z systemem DMM (Defensive Mechanisms Manual) stworzonym przez Phoebe Cramer do pomiaru mechanizmów obronnych. DMM ma trzy skale, które opisują kolejno trzy coraz bardziej dojrzałe mechanizmy obronne: zaprzeczanie, projekcję i identyfikację. System opiera się na spójnej koncepcji teoretycznej nawiązującej do psychoanalizy, lecz wykraczającej poza nią przede wszystkim w tym, że uwzględniony jest rozwój w ciągu życia. System ma też mocne podstawy empiryczne. Dlaczego zatem nie został doceniony przez grupę krytycznych badaczy? Po wnikliwym przeanalizowaniu ich argumentacji [23; por. 64] pojawia się tylko jeden, konieczny wniosek: złe zdanie Lilienfelda, Wooda i Garba [30] na temat DMM wynika z całego szeregu pomyłek i pominięć, które przydarzyły im się podczas studiowania literatury na temat tego systemu. Tytułem przykładu zostanie opisane jedno z badań trafności [11], które Lilienfeld i in. [29] błędnie uznali za badanie rzetelności. Dwie grupy dzieci opowiadały historie do obrazków. W połowie badania informowano je, że w wykonywanym wcześniej zadaniu wypadły dobrze lub źle (po badaniu przeprowadzono staranny *debriefing*). Spodziewano się, że informacja o po-

⁶ Można raczej łączyć ją z poznaniem utajonym, o czym zresztą wspominał sam twórca tego pojęcia, Anthony Greenwald [19].

rażce spowoduje nasilenie mechanizmów obronnych, natomiast informacja o sukcesie nie stanowi do tego powodu, zatem różnica między dwiema grupami – ale tylko w drugiej części badania – będzie widoczna na skalach DMM. Z racji wieku uczestników badania spodziewano się głównie różnic w mechanizmie zaprzeczania oraz projekcji. Wyniki badania dokładnie spełniły te oczekiwania, co stanowi doskonałe potwierdzenie trafności DMM. Tego typu badań Cramer przeprowadziła więcej, a do wyzwania mechanizmów obronnych używała różnego typu zagrożeń psychologicznych, na przykład krytyki układanych przez uczestnika badania opowiadań [10] albo podważania tożsamości płciowej przez rzekome wyniki innych badań [11]. Należy więc uznać, że trafność DMM w mierzeniu poziomu mechanizmów obronnych trzech rodzajów została udowodniona.

Lilienfeld i in. [29] nie widzą jednak tego w ten sposób. W ich przekonaniu opisane wyżej badania były sprawdzaniem stabilności DMM (jednego z rodzajów rzetelności) metodą test – retest. Na tę myśl naprowadził ich zapewne fakt, że porównywano niejako dwa pomiary rozdzielone w czasie. Jednak w ten sposób bada się także trafność, jest to metoda tzw. analizy zmian nieprzypadkowych wyników testu [por. 24]. Sami krytyczni autorzy także musieli chyba dostrzec swoje pomyłki w odniesieniu do systemu Cramer, gdyż w późniejszych swoich publikacjach nie powtarzali już tych zarzutów (choć też ich nie odwołali – po prostu nie wypowiadali się na ten temat).

7. Stosować czy nie stosować? Podsumowanie

Uwzględniając wszystko, co napisano powyżej, można powiedzieć, że mamy oto taką sytuację z TAT, że najbardziej popularne systemy interpretacji – Murraya i Bellaka – nie spełniają standardów, w tym przede wszystkim ich trafność nie została wykazana (a w badaniach wykazywano wręcz nietrafność niektórych popularnych interpretacji). Tych systemów nie powinniśmy więc używać wcale, a już na pewno nie w diagnozie do celów sądowych.

Co w takim razie z systemami, które pomyślnie przeszły konfrontację z danymi empirycznymi – systemem McClellanda, Westena i Cramer? A także z systemem Mursteina do oceny wrogości lub z wieloma innymi systemami do oceny konkretnych zmiennych, o których tu nie wspomniano⁷?

Spójrzmy raz jeszcze na kryteria zamieszczone w rozdziale 2. Pierwsze z nich mogłoby się wydawać spełnione w odniesieniu do wszystkich wymienionych wyżej systemów ugruntowanych empirycznie, gdyby nie wymaganie posiadania norm. Autorzy tych systemów nie opracowali norm ogólnych, zamiast tego zalecali tworzenie lokalnych norm przez poszczególne ośrodki posługu-

jące się danym systemem. W Polsce oznacza to konieczność albo przeprowadzenia badań normalizacyjnych dla populacji polskich, albo wypracowywania norm na własną rękę w różnych placówkach diagnostycznych. Nawet jeśli zostanie zrobione jedno lub drugie, pozostanie kolejna przeszkoda – być może także przejściowa – do posługiwania się empirycznymi systemami interpretacji TAT w diagnozie do celów sądowych. Otóż jak głosi rekomendacja druga, przed rozpoczęciem stosowania danego systemu we własnej praktyce należy przejść trening w jego używaniu. Tymczasem obecnie brak jest w Polsce ośrodków szkolących w zakresie SCORS czy DMM, a system McClellanda jest zarezerwowany dla jednej firmy kontynuującej tradycję zespołu. Być może komuś uda się mimo tych ograniczeń przejść takie szkolenie w kraju lub za granicą. Po rozwiązaniu kwestii norm mogłoby więc stosować na przykład SCORS do oceny skutków nadużyć, DMM do określenia nasilenia i dojrzałości mechanizmów obronnych lub system McClellanda do oceny motywacji związanej z potrzebą osiągnięć, afiliacji i władzy, co może być istotne nie tylko w kontekście kariery zawodowej (pozostają też do wyboru inne, mniej znane skale).

Jednak bezpowrotnie trzeba się rozstać z wyobrażeniem TAT (lub innego narzędzia projekcyjnego) jako „rentgena”, dzięki któremu psycholog prześwietla duszę ludzką i widzi wszystko, co w niej jest ukryte. Techniki projekcyjnymi – ani żadnymi innymi – nie zmierzmy całokształtu psychiki ludzkiej. Zauważyć należy, że systemy bez ugruntowania empirycznego, zwłaszcza system Bellaka, to właśnie usiłowały robić – dostarczać diagnościę wiedzę o całej osobowości, o całej psychopatologii. Tak ambitnych zadań nie jest się w stanie wykonywać w naukowej diagnozie psychologicznej, choć laicy, którym psycholog kojarzyć się może z jasnowidzem lub wróżką, tego być może oczekują (i zarazem obawiają się). Sposobami naukowymi można podejmować zadania o wiele skromniejsze, na przykład określić nasilenie trzech potrzeb lub trzech mechanizmów obronnych. Teoria pozwala wyprowadzić z tego wnioski wykraczające poza same te dane surowe, jednak nie nazbyt daleko. I to jest być może najważniejsza rzecz, którą warto zapamiętać na temat technik projekcyjnych lub w ogóle narzędzi diagnostycznych. Jeśli nie ma akurat dostępu do wiarygodnych opracowań, a chce się wiedzieć, czy przedstawiana psychologowi technika jest wartościową techniką diagnostyczną, należy spojrzeć na to jedno kryterium: jak wiele technika ta obiecuje ujawnić na temat uczestnika badania. Jeśli jest to jedna lub kilka, w wyjątkowych przypadkach więcej ściśle określonych zmiennych, to należy dać tej technice szansę. Natomiast wszelkie techniki obiecujące wgląd w głębokie nieświadome warstwy osobowości i ujawniające jakoby rozliczne i różnorodne prawdy na temat uczestników, są technikami dla czarodziejów, a nie dla psychologów!

⁷ Kilka z takich systemów opisano w innej pracy [63].