

Decision-aiding in the process of psychological assessment

KARL WESTHOFF¹, CARMEN HAGEMEISTER & ANJA STROBEL

Abstract

After some central statements on the assessment process, decision-aiding by Decision-oriented Assessment will be presented. Then rules for deciding about the client's question will be explained. After that the basis of every assessment is listed: the requirements and how to gather them. The behavioral equation of Decision-oriented Assessment will be introduced and its usefulness for the structuring of the assessment process will be shown. One further crucial phase of the assessment process is the selection of sources of information. - The basis of every sound psychological assessment is observation of behavior. Important parts of behavior can only be observed by the subjects themselves. Decision-oriented Interviewing enables assessors to gather these self-observations. The Instrument for the Description of Interviewer Competence in Proficiency Assessment (DIPA, Strobel, 2004) highlights the strengths and weaknesses of an interviewer when gathering self-observations from subjects. - Inevitably in the assessment process all pieces of information must be integrated so that, if agreed, well-founded recommendations can be derived.

Key words: Psychological assessment, decision-aiding, assessment process, task analysis

¹ Prof. Dr. Karl Westhoff, Technische Universität Dresden, Institut für Psychologie II, 01062 Dresden, Germany; email: karl.westhoff@tu-dresden.de

1. Introduction

In each applied field of psychology, sound assessment results are the basis for successful and satisfying decisions. The assessment process is conceived as a sequence of many inevitable decisions under uncertainty. Decision-Oriented Assessment (DOA, Westhoff & Kluck 2003, first published in 1991) is presented here as decision aiding technology for all these decisions.

Although the assessment process is not the focus of psychological research, explicitly understanding it as a sequence of inevitable decisions is very helpful in practice (Fernandez-Ballesteros et al., 2001). The assessment process starts with the decisions the assessor has to make about the client's question (Hagemeister & Westhoff, 2002). - Like Fernandez-Ballesteros et al. (2001), we prefer the more general term *client's question* to the term *referral question*, which is limited to clinical psychology. - Then the assessor needs knowledge about the field in question, which can be compiled in a requirements profile (Westhoff & Kluck, 2003). How can one gather relevant requirements? Lots of possible predictors are at hand in psychological assessment (Meyer et al., 2001). The problem is not to overlook a helpful predictor (Westhoff, 2005). Another problem is the selection of valid sources of information. Observations - and very often only observations made by the assessee him/herself - are the basis of sound psychological assessment. But how can self-observations be collected as validly as possible, for example in an in-depth interview? Assessors - even experienced ones - run the risk of interviewing in a sloppy way. Continuous feedback about their kind of interviewing is one measure to help interviewers (Strobel, 2004). The data collected in the assessment process must be integrated in order to answer the psychological questions into which the client's question was divided. The answers to these psychological questions then have to be integrated into an answer to the client's question (Fernandez-Ballesteros et al., 2001). In these inevitable phases of the assessment process, lots of errors and mistakes can occur. In order to improve the assessment process, Decision-Oriented Assessment provides the assessor with helpful devices, for example checklists, heuristics and collections of rules extracted from empirical literature (Westhoff & Kluck, 2003).

2. Aiding the decision-maker in the assessment process

2.1 Guidelines for the assessment process

In continental Europe psychologists typically start out from a general psychological assessment. In the Anglo-Saxon countries another approach dominates. There, one finds as many forms of psychological "assessment" (ways to assess) as there are fields of applications for psychological assessment and intervention.

However, if we examine the assessment process, it is basically the same in all areas of application in psychology. Naturally there are differences in the forms of questioning and in the instruments and information sources. So it is an advantage to have a working knowledge of the assessment process in its most general form, as this makes it easier to work one's way into any specific field of application.

For this reason, the European Association of Psychological Assessment (EAPA) founded a task force to develop Guidelines for the Assessment Process (GAP). The choice of the

acronym GAP was intentional, because there is indeed a gap between the multitude of research into psychological assessment and, for example, the diverse standards for psychological tests (e.g. American Educational Research Association, 1999) on the one hand, and the all-encompassing assessment process on the other. The task force published these guidelines at first in English (Fernandez-Ballesteros et al., 2001), and then also in German (Westhoff, Hornke & Westmeyer, 2003).

2.2 Neither descriptive nor normative: Decision-aiding

The results of decision research have long been utilised for decision aiding. One approach of decision-aiding, for example, uses checklists to support the decision-makers. Checklists always serve to ensure that decision-makers do not forget anything, something which can happen all too easily during complex decision-making processes.

If we want to examine the assessment process, we can do this scientifically in completely different ways. One approach which is very useful for assessment practice is to list up all the decisions that one necessarily makes, i.e., that one cannot avoid making in an assessment process (Westhoff & Kluck, 2003).

In the case of decisions under uncertainty, decision-aiding does not lead to the "right" decisions. It is typical for such decision-making processes that one can never know in advance what the right choice is. But decisions can be made in such a way that, later, the decision-maker at least will not regret the way the decision was reached.

Decision-makers tend to later regret their behavior during decision-making if, for example, they find out that they did not obtain important information, or did not give such information any, or only insufficient consideration (Janis & Mann, 1977). Not giving adequate consideration to all available alternatives is another possible mistake that can be frequently observed. Even essential aims and values are very often not given sufficient consideration, and this may lead to wrong decisions. Rational decision criteria may also be helpful, especially if they are actually applied consistently (Grove et al., 2000; Sawyer, 1966). It is also very important to define well-founded decision rules before starting with the assessment work. Another point is that assessors often forget the deliberations and predefinitions for the implementation of decisions. This also involves thinking and planning in terms of alternatives.

This all shows clearly that decision-aiding is not descriptive. It does not describe the behavior of decision-makers. Nor does it make any prescriptions. So it is not normative, but rather provides support that the decision-maker may use or not, as the case may be. The most famous decision-aiding approaches are those of Taylor and Russell (1939) and Cronbach and Gleser (1965) which, of course, are integrated in Decision-Oriented Assessment like all other approaches that are helpful for an assessor. If, however, an assessor does not make use of decision-aiding devices, a judge for example may ask him/her why he/she did not adhere to the state of the art as described in the Checklists for Decision-Oriented Assessment.

The rules of Decision-Oriented Assessment were one of the bases for the Guidelines for the Assessment Process (Fernandez-Ballesteros et al, 2001) mentioned above and are therefore in line with these European Guidelines covering the whole assessment process.

2.3 Inevitable and optional phases of the assessment process

The assessment process has the following three inevitable phases: *planning*, *collecting data*, and *combining data*. Additionally, if this has been agreed, there may also be the following three optional phases: *recommendation*, *implementation of the recommendation*, and *evaluation* of the assessment and intervention.

In each phase, a series of crucial decisions must be made. This already starts with the decision whether to deal with a particular question or not, or whether to only deal with the question once a workable formulation of the question has been explicitly agreed with the client.

3. Decision aiding in the assessment process

3.1 Rules for deciding about the client's question

In Decision-Oriented Assessment, the following rules were empirically tested in the education and training of psychology students (Hagemeister & Westhoff, 2002, 2003):

- Rule 1 *"Is the client's question formulated unambiguously?"* - Questionings breaking this rule are very vague, so that it is not clear which problem is to be examined.
- Rule 2 *"Is a psychologist the relevant expert?"* - Sometimes a client's question from another field is brought up to the psychologist; often from the field of medicine.
- Rule 3 *"Does enough knowledge exist to answer the client's question?"* - Questionings breaking this rule ask for a prediction when either a prediction is generally impossible or when the fact is disregarded that predictions can only be made by taking probabilities into account.
- Rule 4 *"Is working on this client's question permissible under the law?"*
- Rule 5 *"Is it ethical to work on this client's question?"* - Here we distinguish two aspects: a) infringements of third persons' rights and b) infringements of the rights of the client. Questionings breaking Rule 5a are brought up by individuals and ask for the examination of a third person who has the right to self-determination in this field. Regarding Rule 5b, we see assessment as a help for the contractor to make decisions, which means that a contractor cannot delegate his/her responsibility for the decision to the psychologist and ask what he/she "should" do.
- Rule 6 *"Does the client's question impose no restrictions which are unnecessary?"* - Questionings breaking this rule only ask for aspects of the examination, e.g. they ask for the advantages and not for the disadvantages or require only some aspects to be examined.
- Rule 7 *"Does the questioning not tell which intervention has to take place?"* - These questionings propose the intervention the psychologist has to take in order to solve a problem.

3.2 Concrete valid requirements

Once the questioning has been agreed with the client, the psychologist has to gather information about what concrete requirements or demands are implied by the question. The more concrete the information on these requirements is, the easier it is to find explanations, to make prognoses or to propose interventions.

A very well-known approach towards elaborating requirements is the Critical Incident Technique (CIT) (Flanagan, 1954). The Critical Incident Technique consists of a "... set of procedures for collecting direct observations of human behaviour in such a way as to facilitate their potential usefulness in solving practical problems and developing broad psychological principles ..." (Flanagan, 1954, p. 327). These direct observations are called Critical Incidents (CIs). A CI consists of (a) the description of a situation which discriminates the task performance of an individual, and (b) the description of the individual behavior shown in that situation. A critical incident is a typical and significant initial situation to which a person has to react. This can be done in various ways, and in a further step of the process experts can evaluate the line of action, classifying it from very good to very bad. Performance-differentiating situations are frequently equivalent to tasks that need to be coped with in a specific job. A complete collection of all typical and significant performance-differentiating situations and the corresponding behavior patterns, described and evaluated as very effective to very ineffective by experts, forms the basis for further action.

The CIT is a partially structured approach requiring intensive reading of the literature, a certain expertise, and even then often results in highly differing approaches to the same topic (e.g. Gremler, 2004). For this reason Koch and Westhoff (2005) presented an approach for the field of personnel assessment, based on the CIT and further developments of the CIT as well as recent promising approaches in task analysis, which they call Task-Analysis-Tools (TAToo). The final aim is to enable even non-specialists to create a well-founded requirements profile based on concrete behavior descriptions.

Following the CIT, all typical and significant performance-differentiating situations are classified according to similarity. Each class of situations is then given a name which describes the behavior, e.g. "presenting specialist matters understandable". In the next step these generalized behavior descriptions are related to psychological concepts and constructs, e.g. here "verbal intelligence". In this second step users of the CIT frequently choose generalized names that are not scientifically investigated concepts or constructs but just formulations in everyday language e.g. "social competence" (Neubauer, 2005, pp. 96-97). In contrast, TAToo advocates drawing on the theoretical and practical resources of over 50 years of intensive and productive psychological research and, as far as possible, only using scientifically established concepts and constructs. This has decisive advantages for the work in hand, but no disadvantages compared to self-invented terms which have not yet been examined scientifically. Furthermore different tasks profiles can be more easily compared on this more abstract level.

3.3 *The behavior equation of decision-oriented assessment*

Behavior is a function of a range of psychological and non-psychological factors and their interaction.

$$\mathbf{B} = \mathbf{f}_1 (\mathbf{En}, \mathbf{O}, \mathbf{C}, \mathbf{Em}, \mathbf{M}, \mathbf{S})$$

Behavior is a function of the following six groups of variables: (1) Environment variables (En), (2) Organism variables (O), (3) Cognitive variables (C), (4) Emotional variables (Em), (5) Motivational variables (M), (6) Social variables (S), and their interactions (subscript I).

This proposed behavioral equation serves several purposes: (1) It groups the variables which are important for the explanation, prediction and modification of individual behavior into a few groups, and therefore allows (2) this otherwise vast area to be structured. (3) It enables the assessor to check whether all important aspects have been included in the psychological assessment process. It is therefore meant to be an aid to practical work. - The six groups of variables can again be grouped into two classes: (1) non-psychological groups of variables (En, O), and (2) psychological groups of variables (C, Em, M, S).

The non-psychological variables in the environment (En) refer to external living conditions such as residential situation and financial situation. The non-psychological variables in the organism group (O) include all physical conditions such as illnesses, disabilities or ability to withstand physical strain. Everybody knows how strongly these variables influence our behavior. Therefore, in our experience, it is understandable for non-psychologists when we deal with these areas separately.

The group of psychological variables called cognitive (C) refers to mental ability and includes, for example, concentration, memory, intelligence and creativity. This is the group of variables which can best be measured with tests. However, it is completely different with the emotional variables (Em). These are often overlooked, or only the person's ability to withstand emotional strain is assessed. However, in certain classes of situation, emotions such as fear, guilt or love are often decisive for what people do. Motivation determines behavior, but the question is which specific motivational variables (M) are really good predictors for the behavior specified in the client's question. The strengths of particular motives can be measured, for example the strength of the performance motive, and these strengths can be used to predict individual behavior. However, in our experience, better predictions can be made if the relevant values, goals, desires, beliefs and expectations of the subject are assessed. Social variables (S) influence human behavior in many ways. In psychological assessment, it is therefore necessary to consider both the binding norms and responsibilities of a person, as well as the influence of "significant others". The network of social relationships within which a person lives can never be left unconsidered when an assessor wants to explain and predict behavior successfully.

3.4 Grouping of behavior-differentiating situations

The behavioral equation can help to select variables under which performance-differentiating situations can be grouped together. Behavior, as we have seen here, is influenced by many factors, and this also applies to behavior in performance-differentiating situations. However, each individual situation and the behavior shown in it usually has one distinct emphasis under one variable.

On the basis of these assigned performance-differentiating situations those variables can be identified which are significant for the explanation, prediction or modification of the behavior. And in doing this an assessor can draw on replicated empirical correlations between a relevant variable and a specific behavior pattern. This is a key advantage over terms formulated purely in everyday language, which allow no recourse to the findings of scientific research, because, obviously, there are not any.

After applying task analysis tools, the psychologist does not only know the performance-differentiating situations in which a candidate must prove him/herself, but the psychologist also knows which scientific concepts and constructs he/she must obtain further information on if he/she wants to be able to explain, predict or influence the behavior of the test person in these situations.

3.5 Selecting sources of information

The psychologist can now consider which information sources are available to him/her in principle for a particular questioning, and which of these he/she wants to use. With the help of the “Checklist Selection of Sources of Information” (Westhoff & Kluck, 2003, p. 245) (Figure 1), he/she can ensure that he/she does not overlook any useful information sources. This is the next step in the development of an assessment strategy.

1. *Standardized procedures*: tests? / questionnaires / work samples? / behavioral observations?
2. *Partly-standardized procedures*: behavioral observations? / decision-oriented interviews?
3. *Non-standardized procedures*: interviews? / certificates? / files? / medical reports? / others?

Figure 1:

“Checklist Selection of Sources of Information” (Westhoff & Kluck, 2003, p. 245)

4. Behavior observation data are the basis

The observation of concrete behavior in a specific class of situations is the basis of every useful piece of assessment information. The behavior can be observed by an external observer or by the subject him/herself. The behavior can be maximum possible behavior, as intended for example in performance tests, or typical i.e. usual behavior, like in the majority of personality questionnaires.

The most significant behavior is usually the most recent behavior. Self-observations can be particularly problematic here, because they may be subject to a range of biases. The relevant sections of concrete behavior in performance-differentiating situations can be recorded by external observers or by devices, where this is possible.

However, many decisive factors for human behavior cannot be observed from the outside or recorded by means of questionnaires. Questionnaires, in their standardized way, merely record a small section of the self-perception of a person. This is where the partially structured in-depth interview for the elicitation of qualitative information gains crucial importance (Qualitative information tells us who behaves how under what conditions). After all personal aims, values, attitudes and beliefs for example determine human behavior to a large extent.

4.1 Decision-Oriented Interviewing

Decision-Oriented Interviewing (DOI) compiles all the valid rules for the collection of the significant pieces of information, and, additionally, like the DOA approach as a whole, DOI is open to any empirical and theoretical progress. Naturally by the DOI-rules selection interview guides can be constructed that result in quantitative data. Yet these kinds of interviews are restricted to a few but important kinds of clients' questions. Most interviews in psychological assessment, however, serve to collect qualitative information regarding an individual case. To illustrate, the DOI rules for the structuring of in-depth interviews (Westhoff & Kluck, 2003, p. 248) are presented below (Figure 2) and a few which are exemplary in their relevance in psychological assessment are explained.

It is helpful to view such interviews as indirect observations of behavior which are planned, realized and summarized according to well-founded and empirically tested rules. In this way, an assessor can collect key information on, for example, the success or failure of a person that no other instrument can give.

Here it is essential that the interviewee can and should give spontaneous descriptions of concrete behavior in performance-differentiating situations. So it should be a relatively easy task for the interviewee. Furthermore, the interview should be planned so that the interviewee will also later be able to observe him/herself in future in performance-differentiating situations with relative ease.

- | |
|--|
| <ul style="list-style-type: none"> – Is the Decision-Oriented Interview planned/executed/analyzed according to the state of the art? – Does the interview guide help the subject to report about concrete behavior and experiences? – Does the interview guide enable the subject to report observations spontaneously? – Can the subject use the interview guide as a basis for own observations? |
|--|

Figure 2:

Example items from the "Checklist planning a Decision-Oriented Interview (DOI)"
(Westhoff & Kluck, 2003, p. 248)

It is a great help to the interviewee if the questioning is clarified one last time before the interview begins, meaning that assessor and subject together review their collective aims once more and explicitly confirm their agreement on the proposed line of action. To ensure that the interviewee knows whether the assessor has understood him/her in all essential points, at the end of each section the assessor gives a summary of the relevant observations reported by the interviewee. This also has a highly motivating effect on the subject because it lets him/her see that the assessor is interested in him/her and that the information given has been correctly received. Here, the assessor can use the "*Checklist Broad Structure of an Interview Guide*" (Westhoff & Kluck, 2003, p. 248)

In the "*Checklist Detailed Structure of an Interview Guide*" (Westhoff & Kluck, 2003, p. 248), the rules are compiled for the detailed formulation and structuring of an interview guide. We look in detail at just two of them here, both of which are clearly of special relevance for psychological assessment. The first rule is: Every question should relate to observable behavior, meaning behavior visible from the outside, or feelings and thoughts which precede, accompany or follow this behavior. Before a decisive exercise, an athlete may for example think of something inappropriate like, for instance, "I'm going to win". Such a thought has nothing to do with the activity to be carried out. The athlete then loses concentration because his/her mind is on something else and not on what he/she wants to do in this moment. In contrast, it may be helpful, for example, if he/she recapitulates the correct sequence of actions to be carried out. Mistakes or weaknesses in the realization of a specific behavior may therefore be based on wrong thoughts. - The second rule is: The interviewer should only ask effective questions. In order to do this he/she should stick to the rules in the checklist regarding effective questions.

In order to use only effective questions an interviewer can regard the rules compiled in the "*Checklist for Formulating Effective Questions*" (Westhoff & Kluck, 2003, p. 249). It is important, for example, that subjects are asked to report what they felt, thought or did in a concrete situation. They should not describe what "one" does in such situations. Furthermore, they should not replace their observations with explanations of their behavior because lay persons' attempts to explain behavior can easily be wrong.

An old rule of interviewing is that a single question should always only ask about a single thing, i.e. the single idea rule (Cannell & Kahn, 1954). Many interviewers ignore this rule. The result is that interviewees only report on a part of what they were asked about.

Interviewers must be given special training in formulating questions as short and at the same time as precisely as possible. Long questions usually contain an unnecessary large number of nouns. Every noun is an abstract term which requires the interviewee to think about how exactly it is meant in this context. Questions which use verbs rather than nouns are better and easier to understand.

Most interviewers do not notice when their questions are leading questions, i.e. suggestive. Yet closed questions which can be answered with a single word like, for example, yes or no are in fact potentially suggestive. Multiple choice questions are not so good either because they do not allow the interviewee to respond freely. In their questions, interviewers often assume as given something which may actually have been different. The question "How many laps do you run to warm up?" is a leading question because an athlete can also warm up in other ways.

In normal daily life, when we observe processes, we ask about the whys and wherefores, or what the reasons are. But human behavior is not determined by individual "reasons". In

each case there is a combination of many factors at work, unless it happens to be a reflex. If we ask interviewees for the reasons for their behavior, what we get is not "the" motive, as is often incorrectly assumed, but merely causal attributions. We would probably get useless information with the question "Why is your training good for you?" Causal attributions are however only a relatively insignificant part of many motivational variables.

The most important motivational variables are those relating to learning psychology, namely reinforcements and punishments. After a reinforcement, the frequency of a behavior increases; after a punishment, it decreases. We can collect these data with questions like "What do you find positive about your training?" and "What do you find less positive about your training?" It would be inappropriate to ask directly about reinforcements and punishments, as non-specialists would not know what the terms mean.

So how can an interviewer in proficiency assessment find out whether and to what extent he/she is adhering to rules which have been proved empirically as "good"? The answer is the Instrument for the Description of Interviewer Competence in Proficiency Assessment (DIPA, Strobel, 2004).

4.2 The Instrument for the Description of Interviewer competence in Proficiency Assessment (DIPA)

In spite of extensive findings on and recommendations for the structuring of assessment interviews, there are still serious shortcomings in the transfer to practice. The interviewer him/herself has a crucial influence on the quality of the interview. Especially the aspect of experience, which is generally considered to be essential for quality, has often proved problematic (see Dipboye & Jackson, 1999, for review). Mere theoretical knowledge of the effects on the interview, or a single training are not enough to ensure the quality of the interview process. What is missing is a continuous feedback for the interviewer that can be implemented in daily routine without excessive effort. This is where the Instrument for the Description of Interviewer Competence in Proficiency Assessment (DIPA) comes into play.

The DIPA was developed on the basis of Decision-Oriented Interviewing (Westhoff & Kluck, 2003) and its extension Requirements Profile for Psychological Assessment Interviews (Kici & Westhoff, 2000, 2004). Here, the contents of the requirements profile, which was mainly intended for psychological assessment interviews in general, were adapted to the special features of the proficiency assessment interview. To increase practicability when used in the human resources sector, various aspects were shortened, especially those relating to the written account of assessment results in the report. In addition, special consideration was given to current findings of interview research (cf. Strobel, 2004) as well as recommendations and directions by experienced practitioners. Special attention was paid to the economic and practical aspects of the instrument. The scope and along with it the corresponding time needed for the whole process were tailored to the typical lack of time prevalent in business, but without dispensing with a differentiated description of the essential aspects.

The DIPA consists of two variants of equivalent content – one for the evaluation of a single interview and one for the overall evaluation of several interviews conducted by an interviewer. Each of these variants consists in turn of the three parts, planning, realization and summarization of the single interview or a series of interviews. For the evaluation of interviewer behavior in the three areas, the user has 145 items on 6 pages at his/her disposal.

Interview Guide					
7	What is the structure of the interview guide?	☐ in keywords	☐ mainly in keywords	☐ mainly formulated questions	☐ formulated questions
8	Is the structure of the interview guide based on a requirement profile?			☐ yes	☐ no
9	Does the interview guide contain a part for a formal welcome?			☐ yes	☐ no
10	Does the interview guide contain a part for the collection of information?			☐ yes	☐ no
Formulation of the questions					
Are the questions...					
73	• based on concrete behavior of the applicant?	☐ always	☐ often	☐ seldom	☐ never
74	• as short as possible?	☐ always	☐ often	☐ seldom	☐ never
75	• adequately open?	☐ always	☐ often	☐ seldom	☐ never
76	• adequately straight?	☐ always	☐ often	☐ seldom	☐ never

Figure 3:
Detail from the DIPA

To make it easier to use, the majority of the questions were structured so that a uniform response format could be applied. One response format is affirmation vs. rejection (yes – no) and the other format is frequency rating (always – often – seldom – never). Figure 3 shows as an example a detail from the DIPA.

The DIPA thus facilitates differentiated, concrete and behavior-descriptive feedback on interviewer behavior during planning, realization and summarization of assessment interviews. A scoring key – developed via a multi-stage survey among experts – has a rating scale for each item listed in the instrument from 0 (unacceptable), 1 (problematic), 2 (acceptable) to 3 (optimal). This enables the user by means of color-coded evaluation transparencies to get a survey of his/her own strengths and weaknesses in interviewing.

Empirical testing and acceptance in practice

The DIPA was tested in seven studies, with successive improvements being added on the basis of the findings and the feedback from the participants. The test participants were practicing assessors and students with psychology as their main subject. Furthermore, the above mentioned scoring key was compiled with the help of expert judgments. In the last two studies 51 practicing assessors and 80 students with psychology as their main subject evaluated an interview guide, the realization of an interview in a transcript-version and a summary of the interview (interview notes and a short report with ratings) on the basis of the DIPA-items. Four weeks later, the students (N=63) got the same task to assess the retest-reliability of the instrument. After test and retest the participants had to fill in a final questionnaire with items on economical and practical aspects of the DIPA. The inter-rater-reliability was assessed using the kappa-coefficient for more than two assessors (Fleiss, 1971) or the “weighted-kappa-coefficient” (Cohen, 1968), respectively. To assess retest-reliability, Spearman-rank-correlations were calculated. The validity of the DIPA

was assessed via percent agreement with a master solution. The master solution contained the intended results for the ratings, that is the correct answers for each item as they arose from the construction of the material.

The measured kappa-coefficients for objectivity were in the mid range (.26 - .53), with rough agreement among the participants of over 70% in each case. The coefficients for re-test-reliability were at over .50 for all parts of the DIPA. Validity, rated via percent agreement with the master solution, was satisfactory at 58 to 86 percent, with mid scores for the evaluation of the very complex realization and summarizing parts (58 to 65 percent) of the interview and relatively high scores for the planning part (79 to 86 percent) of the DIPA. The economic and practical aspects of the DIPA were widely assessed as positive. The majority of the participants rated the items as intelligible and practical. While two thirds of the practitioners and one third of the students assessed the DIPA as too large, at least 60 percent of the practitioners and over 90 percent of the students stated that they could imagine using the DIPA in their professional practice and also felt that it would be useful. Recent studies with both, students and practitioners, confirm the high acceptance and practicability of the instrument in training and daily work.

All in all, the psychometric features of the DIPA show that the instrument facilitates an objective, reliable and valid assessment of interviewer competence (Strobel, Plath & Westhoff, 2003; Strobel & Westhoff, 2004; summarized in Strobel, 2004). This is an instrument which meets scientific criteria in terms of its quality and validity, and provides practitioners involved in aptitude assessment with an efficient and productive tool.

5. Combining all information and recommendations

In addition to planning and realization of assessment measures, the evaluation of different sources of information has a special relevance (Eckert & Westhoff, 2002). Here, too, practitioners get support from the set of guidelines of Decision-Oriented Assessment, as they do when it comes to combining all information and developing recommendations. In these activities it is essential that the assessor arrives at the results of the assessment process according to pre-defined and well-founded rules. The recommendations to be made should then be based, with total comprehensibility, on the assessment results.

Finally, from the "*Checklist Recommendations and Suggestions in a Psychological Report*" (Westhoff & Kluck, 2003, p. 251) a few central rules will be presented which should be noted when the psychologist has to make recommendations, that is he/she is asked to aid the client in his/her decisions.

Firstly, it is important to describe all possible alternatives (Janis & Mann, 1977, p. 11). Advisors often restrict themselves to just a few alternatives, and not necessarily always the best ones. Recommendations in themselves do not usually help much, unless the conditions for their implementation are mapped out in detail. There must be explicit agreement on who will do what, when and how. The achievable goals must be specified, as must the potential, perhaps unpleasant consequences of each possible course of action. Only if psychological assessors proceed like this can they support decision-makers in such a way that these can make their own decisions. Only if the assessor states all facts and all considerations explicitly can a decision-maker get such comprehensive advice that he/she does not later regret the path he/she took to a particular decision.

6. Discussion

If a person needs comprehensive psychological advice, each decision made by the person must be reached with painstaking care. It is essential that nothing important is overlooked. The checklists of Decision-Oriented Assessment are an aid here.

For the systematization of necessary decision-making, existing knowledge in psychology on significant variables can be structured according to the behavior equation of Decision-Oriented Assessment.

A valid requirements profile is the indispensable starting point for any form of assessment and intervention. Such a requirements profile must be developed for each specific questioning. The requirements profile can be drawn up using the task analysis tools (TAToo). Here, the expert knowledge is gathered and systematically organized according to explicit rules. This then forms the empirical basis for every form of assessment and intervention.

The majority of the individual pieces of information on a person are acquired through observation and recording of individual behavior. Behavior which can be observed externally is relatively easy to record with maximal objectivity. The internal behavior like emotions and thoughts, however, can only be observed by the subject him/herself. And these emotions and thoughts are often crucial.

Here, standardized procedures like questionnaires are often of little help, as it is the individual behavior and experience of the individual that really matters. Decision-oriented interviewing (DOI) provides the rules and procedures required for gathering valid information in in-depth interviews.

In order to combine all assessment results in an optimal way the rules of Decision-Oriented Assessment can be used. However, these rules are *general* rules. If the combination of information is to be further optimized, question-specific combination rules for frequently occurring questions can be developed and tested. The formulation of recommendations for possible interventions can also be optimized using the respective rules of Decision-Oriented Assessment. And, finally, Decision-Oriented Assessment is also suitable for the evaluation of interventions.

From this outline here, it becomes clear that Decision-Oriented Assessment can be extremely useful for optimizing the whole assessment process including intervention and evaluation.

References

- American Educational Research Association (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Cannell, C.F., & Kahn, R.L. (1968). Interviewing. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (Vol. II., 2nd ed.). Addison-Wesley, Reading, Mass.
- Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision scaled disagreement or partial credit. *Psychological Bulletin*, 70, 213-220.
- Cronbach, I.J., & Gleser, G.C. (1965). *Psychological tests and personnel decisions* (2nd ed.). Urbana: University of Illinois Press.

- Dipboye, R.L., & Jackson, S. (1999). Interviewer experience and expertise effects. In R.W. Eder & M.M. Harris (Eds.), *The Employment Interview Handbook* (pp. 259-278). Thousand Oaks, CA: Sage.
- Eckert, H., & Westhoff, K. (2002). Effects of systematic training in psychological assessment on competence in writing reports. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 23, 407-416.
- Fernandez-Ballesteros, R., De Bruyn, E.E.J., Godoy, A., Hornke, L.F., Ter Laak, J., Vizcarro, C., Westhoff, K., Westmeyer, H., & Zaccagnini, J.L. (2001). Guidelines for the Assessment Process (GAP): A proposal for discussion. *European Journal of Psychological Assessment*, 17, 187-200.
- Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, 51, 327-358.
- Fleiss, J.L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76, 378-382.
- Gremler, D. D. (2004). The critical incident technique in service research. *Journal of Service Research*, 7, 65-89.
- Grove, W.M., Zald, D.H., Lebow, B.S., Snitz, B.E., & Nelson, C. (2000). Clinical versus mechanical prediction: A Meta-Analysis. *Psychological Assessment*, 12, 19-30.
- Hagemeister, C., & Westhoff, K. (2003). Diagnostische Strategien lehren und lernen: Beurteilung diagnostischer Fragestellungen durch Studierende [Teaching and learning assessment strategies: Students' evaluations of clients' questions]. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 24, 97-104.
- Hagemeister, C., & Westhoff, K. (2002). Teaching and learning psychological assessment: Aspects of the client's question. *European Journal of Psychological Assessment*, 18, 252-258.
- Janis, I.L., & Mann, L. (1977). *Decision making – A psychological analysis of conflict, choice, and commitment*. New York: The Free Press.
- Kanning, U.P. (1999). *Die Psychologie der Personenbeurteilung* [The psychology of person evaluation]. Göttingen: Hogrefe.
- Kici, G., & Westhoff, K. (2000). Anforderungen an psychologisch-diagnostische Interviews in der Praxis [Requirements of interviews in practice in psychological assessment]. *Report Psychologie*, 7, 428-436.
- Kici, G., & Westhoff, K. (2004). Evaluation of requirements for the assessment and construction of interview guides in psychological assessment. *European Journal of Psychological Assessment*, 20, 83-98.
- Koch, A., & Westhoff, K. (2005). Konzeption einer Technologie zur Analyse wechselnder Anforderungen an Menschen. Vortrag zum 3. Workshop „Neue Technologien - Ausblick auf eine wehrtechnische Zukunft“ [Concept of a technology for analyzing changing tasks. Lecture given at the 3rd. Workshop: New Technologies – Outlook on Future Military Technology], Bonn.
- Meyer, G.J., Finn, S.E., Eyde, L.D., Kay, G.G., Moreland, K.L., Dies, R.R., Eisman, E.J., Kubiszyn, T.W., & Reed, G. M. (2001). Psychological testing and psychological assessment – A review of evidence and issues. *American Psychologist*, 56, 128-165.
- Neubauer, R. (2005). Wege und Irrwege bei der Definition von Anforderungen [Right and wrong ways to define requirements]. In K. Sünderhauf, S. Stumpf & S. Höft (Eds.), *Assessment Center – von der Auftragsklärung bis zur Qualitätssicherung. Ein Handbuch von Praktikern für Praktiker* [Assessment Center – from contract to quality assurance. A handbook of practice for practitioners] (pp. 89-106). Lengerich: Pabst.
- Sawyer, J. (1966). Measurement and prediction, clinical and statistical. *Psychological Bulletin*, 66, 178-200.

- Strobel, A. (2004). *Das Diagnoseinstrument zur Erfassung der Interviewerkompetenz in der Personalauswahl (DIPA) – Entwicklung, empirische Prüfungen und Akzeptanz in der Praxis*. [Instrument for the description of interviewer competence in personnel selection] Online: <http://nbn-resolving.de/urn:nbn:de:swb:14-1113897309086-89129>.
- Strobel, A., Plath, S.C., & Westhoff, K. (2003). Interviewerkompetenz in Personalauswahl und -entwicklung: Ein Materialvergleich zur Erstellung von Ausbildungsunterlagen [Interviewer competence in personnel selection and development: Comparing materials for training]. *Wirtschaftspsychologie*, 5, 198-201.
- Strobel, A., & Westhoff, K. (2004). Diagnose von Interviewerkompetenz in der Personalauswahl: Ein Instrument für die Praxis [Assessment of interviewers' competence in personnel selection]. In W. Bungard, B. Koop & C. Liebig (Eds.), *Psychologie und Wirtschaft leben – Aktuelle Themen der Wirtschaftspsychologie in Forschung und Praxis* [Living psychology and economics. Current scientific and practical topics of psychology in work and organization] (pp. 172-176). München: Hampp.
- Taylor, H.C., & Russell, J.F. (1939). The relationship of validity coefficients to the practical effectiveness of tests in selection: Discussion and tables. *Journal of Applied Psychology*, 23, 565-578.
- Westhoff, K. (2005). Konstrukte und Operationalisierungen [Constructs and operationalizations]. In K. Westhoff, L.J. Hellfritsch, L.F. Hornke, K.D. Kubinger, F. Lang, H. Moosbrugger, A. Püschel, G. Reimann (Eds.). *Grundwissen für die berufsbezogene Eignungsbeurteilung nach DIN 33430*. 2., überarbeitete Auflage. [Basic knowledge for proficiency assessment according to DIN 33430]. 2nd revised edition (pp. 128-142). Lengerich: Pabst.
- Westhoff, K., Hagemester, C., & Eckert, H. (2002). On the objectivity of oral examinations in psychology. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 23, 149-157.
- Westhoff, K., Hornke, L.F., & Westmeyer, H. (2003). Richtlinien für den diagnostischen Prozess - Zur Diskussion gestellt. [Deutsche Adaptation von: Fernandez-Ballesteros, R., De Bruyn, E.E.J., Godoy, A., Hornke, L.F., Ter Laak, J., Vizcarro, C., Westhoff, K., Westmeyer, H., & Zaccagnini, J.L. (2001). Guidelines for the Assessment Process (GAP): A proposal for discussion. *European Journal of Psychological Assessment*, 17, 187-200.] *Report Psychologie*, 28, 504-517.
- Westhoff, K., & Kluck, M.-L. (2003). *Psychologische Gutachten schreiben und beurteilen*. Vierte, vollständig überarbeitete und erweiterte Auflage [How to write and to evaluate psychological reports. 4th ed.]. Berlin: Springer.