



B

BEHAVIOURAL ASSESSMENT TECHNIQUES

INTRODUCTION

A major impetus for behaviour therapy was disenchantment with the medical model of psychopathology that views problem behaviours as the result of an underlying illness or pathology. Behaviourists assert that both ‘disordered’ and ‘non-disordered’ behaviour can be explained using a common set of principles describing classical and operant conditioning.

Behaviourists believe that behaviours are best understood in terms of their function. Two ‘symptoms’ may differ in form, while being similar in function. For example, Jacobson (1992) describes topographically diverse behaviours such as walking away or keeping busy that all function to create distance between a client and his partner. Conversely, topographically similar behaviours may serve different functions. For example, tantrums may serve to elicit attention from adults or may be an indication that the present task is too demanding (Carr & Durand, 1985). Behaviour therapists try to understand not only the form but also the function of problem behaviours within the client’s environment.

The initial goals of assessment are to identify and construct a case formulation of the client’s difficulties that will guide the clinician and patient towards potentially effective interventions. For the behaviour therapist, this involves

identifying problem behaviours, stimuli that are present when the target behaviours occur, associated consequences, and organism variables including learning history and physiological variables (Goldfried & Sprafkin, 1976). The results of this functional analysis are used to design a behavioural intervention that is tailored to the individual client and conceptually linked to basic learning principles.

Assessing Target Behaviours

The process of defining and measuring target behaviours is essential to behavioural assessment. Vague complaints must be expressed as specific quantifiable behaviours. For instance, anger might include responses such as hitting walls, refusing to talk or other specific behaviours. The client’s goals must be defined in terms of those specific behavioural changes that would occur if treatment were effective.

Target behaviour selection can be complicated by the complexity with which many responses are expressed. Behaviourists have long recognized that many clinical problems involve responses that cannot be readily observed. Some responses such as intrusive thoughts or aversive mood states are private by nature. Others, such as sexual responses, may be private and unobservable due to social convention. Many clinical

complaints may include both observable and private responses. For example, depressed mood and suicidal ideation might be accompanied by crying, or other overt behaviours. Public and private responses may not always appear consistent. For example, an agoraphobic client may enter a shopping mall during an assessment but may do so only with extreme subjective distress.

Cone (1978) suggested that the bioinformational theory of emotion developed by Lang (1971) is useful for conceptualizing clinical problems. Lang (1971) asserted that emotional responses occur in three separate but loosely coupled response systems. These are the cognitive/linguistic, overt behavioural, and psychophysiological systems. A given response such as a panic attack may be divided into physiological responses such as increased heart rate and respiration, cognitive responses such as thoughts about dying or passing out, and overt behavioural responses such as escape from the situation, sitting down, or leaning against a wall for support. Ideally, each response mode should be assessed, there being no a priori reason to value one modality over another (Lang, 1971). Discrepancies are best considered with regard to the particular client, the goals of therapy, and ethical considerations. For example, it may be wise to take verbal reports of pain seriously even if they do not match evidence of tissue damage or physiological arousal.

The triple response conceptualization of clinical problems has encouraged the development and utilization of methods that more or less directly assess each response mode. Overt behaviours have been assessed by direct observation, with psychophysiological assessment used to assess bodily responses, and self-report measures developed to quantify subjective experiences. The apparent link between assessment methods and particular response modes is not absolute. For example, a client might verbally report sensations such as heart pounding, muscle tension, or other noticeable physical changes. However, in some cases, the method of assessment is more closely bound to a particular response mode. This is true of physiological processes such as blood pressure that are outside of the client's awareness, and in the case of thoughts or subjective states that can only be assessed by verbal report. In the following sections self-report measures, direct

observation, and psychophysiological measurements are described in more detail.

SELF-REPORT METHODS

There are several formats for collecting self-report data. These include interviews, questionnaires and inventories, rating scales, think-aloud, and thought-sampling procedures. It is most often the case that an assessment would include several of these methods.

Interviews

The clinical interview is the most widely used method of clinical assessment (Watkins, Campbell, Nieberding, & Hallmark, 1995), and is particularly advantageous in the early stages of assessment. The most salient of its advantages is flexibility. The typical interview begins with broad-based inquiry regarding the client's functioning. As the interview progresses, it becomes more focused on specific problems and potential controlling variables. Interviewing also provides an opportunity to directly observe the client's behaviour, and to begin developing a therapeutic relationship.

The clinical interview also has important disadvantages. Interviews elicit information from memory that can be subject to errors, omissions, or distortions. Additionally, the interview often relies heavily on the clinician to make subjective judgements in selecting those issues that warrant further assessment or inquiry. One could reasonably expect that different clinicians could emerge from a clinical interview with very different conceptualizations of the client (Hay, Hay, Angle, & Nelson, 1979).

Structured and semi-structured interviews were developed in order to facilitate consistency across interviewers. Structured interviews are designed for administration by nonclinicians such as research assistants in large-scale studies. A structured interview follows a strict format that specifies the order and exact wording of questions. Semi-structured interviews are more frequently used by trained clinicians. They provide a more flexible framework for the course of the interview while providing enough structure to promote consistency across administrations. While specific questions may be

provided, the interviewer is free pursue additional information when this seems appropriate. In general, the goal of enhanced reliability has been attained with the use of structured and semi-structured interviews (Matarazzo, 1983). However, the majority of these interviews are designed for purposes of diagnosis rather than more particular target behaviours or functional assessment.

Just as the clinical interview proceeds from a general inquiry to more focused assessment of behavioural targets, other self-report measures vary in the degree to which they assess general areas of functioning versus particular problem behaviours. In general, those measures that assess general constructs such as depression or general domains of functioning are developed using group data and are meant to be applicable to a wide range of clients. Examples of these nomothetic measures include personality inventories and standardized questionnaires. Other self-report methods can be tailored more toward individual clients and particular problem responses. These include rating scales and think-aloud procedures. Each of these methods is described briefly in the following sections.

Questionnaires

Questionnaires are probably the next most common assessment tool after interviews (Watkins, Campell, Nieberding, & Hallmark, 1995). Questionnaires can be easily and economically administered. They are easily quantified and the scores can be compared across time to evaluate treatment effects. Finally, normative data is available for many questionnaires so that a given client's score can be referenced to a general population.

There has been a rapid proliferation of questionnaires over the last few decades (Froyd & Lambert, 1989). Some questionnaires focus on stimulus situations provoking the problem behaviour, such as anxiety provoking situations. Other questionnaires focus on particular responses or on positive or negative consequences. The process of choosing questionnaires from those that are available can be daunting. Fischer and Corcoran (1994) have compiled a collection of published questionnaires accompanied by summaries of their psychometric properties.

Many behaviourists have expressed concern with the apparent reliance on questionnaires both in clinical and in research settings. These criticisms stem in part from repeated observations that individuals evidence very limited ability to identify those variables that influence their behaviour. Additionally, behaviourists point out that we tend to reify the constructs that we measure. This may lead to a focus on underlying dispositions or traits in explaining behaviour rather than a thorough investigation of environmental factors and the individual's learning history. Behaviourists do make use of questionnaires but tend to regard them as measures of behavioural responses that tend to covary rather than underlying traits or dispositions.

Rating Scales and Self-Ratings

Rating scales can be constructed to measure a wide range of responses. They are often incorporated into questionnaires or interviews. For example, a client may be asked to rate feelings of hopelessness over the past week on a scale of 0–8. Clinicians might also make ratings of the client's noticeable behaviour during the interview or the client's apparent level of functioning.

The main advantage of rating scales is their flexibility. They can be used to assess problem behaviours for which questionnaires are not available. Additionally, rating scales can be administered repeatedly with greater ease than questionnaires. For example, rather than pausing to complete an anxiety questionnaire, a client might provide periodic self-ratings of discomfort during an anxiety-provoking situation. The main disadvantage of rating scales is the lack of normative data.

Thought Listing and Think-Aloud Procedures

Clinicians are sometimes interested in the particular thoughts that are experienced by a client in a situation such as a phobic exposure or role-play. The use of questionnaires may interfere with the situation and may not capture the more idiosyncratic thoughts of a particular client. Think-aloud and thought sampling procedures may be used under these circumstances. These

procedures require the client to verbalize thoughts as they occur in the assessment situation. Thoughts can be reported continually in a think-aloud format or the client may periodically be prompted to report the most recently occurring thoughts in a thought-sampling procedure. When the requirements of think-aloud procedures may interfere with the client's ability to remain engaged in the assessment situation, the client may be asked to list those thoughts that are recalled at the end of the task. These procedures carry the advantage of being highly flexible. Like other highly individualized methods, they also carry the disadvantage of lacking norms.

DIRECT OBSERVATION AND SELF-MONITORING

One of the most direct forms of assessment is observation by trained observers. Direct observation can be conducted by clinicians, professional staff, or by participant observers who already have contact with the client. Rather than reporting in retrospect, observers can record all instances of the target behaviour that they witness thereby producing a frequency count. Depending on the type of target response, this task could be arduous. Recording all instances of highly frequent and repetitive behaviours can place undue demands on observers. There are several ways to decrease the demands on the observer and thereby facilitate more faithful data collection. One option is the use of brief observation periods. For example, a parent might be asked to record the frequency of the target behaviour at intervals during those specific situations when the behaviour is probable. When the target behaviour is an ongoing response, the observer might employ momentary sampling procedures and periodically check to see if the behaviour is occurring. For a more thorough discussion of alternative procedures of direct observation see Baird and Nelson-Gray (1999).

Direct observation carries some disadvantages. It can be costly and time-consuming. In the strictest sense it would be favourable to utilize multiple observers so that the concordance of their recording could be checked. It has been shown that the reliability of observations is enhanced when observers know that the data will

be checked (Weinrott & Jones, 1984). However, this may not be practical, particularly in clinical settings. The use of participant observers may be a less costly alternative in many cases.

Direct observation can also result in reactive effects. Reactivity refers to changes in behaviour that result from the assessment procedure. Making clients aware that they are being observed can alter the frequency or form of the target response (Kazdin, 1979). This can occur even with the use of participant observers (Hay, Nelson, & Hay, 1980). The variables that influence observee reactivity are not well understood. For ethical reasons, it may be unwise to conduct observations without the client's awareness.

Self-Monitoring

In self-monitoring procedures, the client is asked to act as his or her own observer and to record information regarding target behaviours as they occur. Self-monitoring can be regarded as a self-report procedure with some benefits similar to direct observation. Because target behaviours are recorded as they occur, self-monitored data may be less susceptible to memory related errors. Like other self-report methods, self-monitoring can be used to assess private responses that are not amenable to observation. Self-monitored data also have the potential to be more complete than that obtained from observers, because the self-monitor can potentially observe all occurrences of target behaviours (Kazdin, 1974).

There are several formats for self-monitoring. Early in assessment, a diary format is common. This allows the client to record any potentially important behaviours and their environmental context in the form of a narrative. As particular target behaviours are identified, the client may utilize data collection sheets for recording more specific behavioural targets and situational variables. When behaviours are highly frequent or occur with prolonged duration, the client may be asked to estimate the number of occurrences at particular intervals or the amount of time engaged in the target response.

It is often desirable to check the integrity of self-monitored data. Making the client aware that their self-monitored data will be checked is known to enhance the accuracy of data collection (Lipinski & Nelson, 1974). Self-monitored data

can be checked against data obtained from external observers or can be compared to measured by-products of the target response. For example, self-monitored alcohol consumption can be compared to randomly tested blood alcohol levels.

Among the disadvantages of self-monitoring are its demands on the client for data collection and the lack of available norms. Like direct observation, self-monitoring also produces reactive effects. However, this disadvantage in terms of measurement can be advantageous in terms of treatment. This is because reactive effects tend to occur in the therapeutic direction, with desirable behaviours becoming more frequent and undesired behaviours tending to decrease. This temporary effect of the procedure can produce some relief for the client and help to maintain an investment in treatment. More information on self-monitoring methods is provided in the self-monitoring entry in this volume.

PSYCHOPHYSIOLOGICAL ASSESSMENT

Psychophysiological assessment is a highly direct form of measurement that involves assessing the byproducts of physiological processes that are associated with behavioural responses. For instance, a cardiometer can be used to measure electrical changes associated with activity of the heart. While clients can verbally report many physiological changes, a direct measurement via instrumentation carries several advantages. Physiological measures can be sensitive to subtle changes and to physiological processes that occur without the client's awareness. They can also provide both discrete and continuous data with regard to physiological processes while requiring only passive participation from the client (Iacono, 1991). Additionally, most clients lack familiarity with psychophysiological measurement making deliberate distortion of responses improbable (Iacono, 1991). Korotitsch and Nelson-Gray (1999) provide a more detailed discussion of psychophysiological measures.

The main disadvantage of psychophysiological measurement is the cost of equipment and training. This problem is compounded by the observation that it is often desirable to include

measures of multiple physiological channels. For example, there can be substantial variance across individuals in the degree of response exhibited on a given physiological index. Those measures that are most sensitive for a given individual may not be included in a limited psychophysiological assessment. With technological advances in this area, less costly instrumentation will likely become more available.

FUTURE PERSPECTIVES

Over the past two decades, research devoted to direct observation and self-monitoring procedures has declined dramatically. This trend has been mirrored by a rapid proliferation of questionnaires (Froyd & Lambert, 1989) and research examining their psychometric properties. One likely reason for this shift is the current climate of managed healthcare. The goal of more efficient and less costly healthcare has created pressure for more rapid and inexpensive forms of assessment and treatment. Psychophysiological recording equipment is simply too expensive for most clinicians to afford and maintain. The task of training and paying trained observers can also be costly. Even when participant observers are used, the procedure can place inordinate demands on these individuals. While self-monitoring is less costly, it does place more demands on the client and more time is required to obtain useful information beyond an initial interview. In general, the more direct methods of behavioural assessment have the disadvantage of also being more costly and time consuming. The trend toward more rapid assessment seems to select for brief, easily administered, and relatively inexpensive questionnaires and rating scales. There have been calls for more research devoted to behavioural assessment methods (Korotitsch & Nelson Gray, 1999; Taylor, 1999). This research might lead to more efficient methods for implementing these assessment procedures. There is also a need to determine if the data from these assessments facilitates more efficient and/or effective treatment (Korotitsch & Nelson-Gray, 1999). If empirical support for the utility of behavioural assessment techniques is generated, this may help to increase the receptiveness of third party payers to the use of these procedures.

CONCLUSIONS

The goals and conduct of behavioural assessment are directly linked to learning theory and to the goal of altering behaviour through the use of behavioural principles. The hallmark of behavioural assessment is an emphasis on the function rather than the form of problem behaviours, and on the specification of problem behaviours, as well as their environmental and organismic controlling variables in more detail than is typical of diagnostic classification. While diagnostic assessment tools might be included, behavioural assessment demands further molecular analysis of specific target behaviours and controlling variables.

Behaviour therapists have long recognized that clinical problems are often part of the client's private experience, and that many are a combination of verbal, physiological, and overt behavioural responses. A comprehensive assessment considers each of these modalities. While these ideas are still fundamental in behavioural assessment, the more costly and time-demanding methods behavioural assessment are becoming more difficult to include in clinical assessment and are less apt to be the focus of research.

References

- Baird, S. & Nelson-Gray, R.O. (1999). Direct observation and self-monitoring. In Hayes, S.C., Barlow, D.H. & Nelson-Gray, R.O. *The Scientist-Practitioner* (2nd ed., pp. 353–386). New York: Allyn & Bacon.
- Carr, E.G. & Durand, V.M. (1985). The social-communicative basis of severe behaviour problems in children. In Reiss S. & Bootzin, R.R. (Eds.), *Theoretical Issues in Behaviour Therapy* (pp. 220–254). Orlando: Academic Press.
- Cone, J.D. (1978). The behavioural assessment grid (BAG): a conceptual framework and a taxonomy. *Behaviour Therapy*, 9, 882–888.
- Goldfried, M.R. & Sprafkin, J.H. (1976). Behavioural personality assessment. In Spence, J.T., Carson, R.C. & Thibnut, J.W. (Eds.), *Behavioural Approaches to Therapy* (pp. 295–321). Morristown, NJ: General Learning press.
- Fischer, J. & Corcoran, K. (1994). *Measures for Clinical Practice: A Sourcebook*. (2 vols., 2nd ed.). New York: Macmillan.
- Froyd, J.E., Lambert, M.J. & Froyd, J.D. (1996). A review of practices of psychotherapy outcome measurement. *Journal of Mental Health*, 5, 11–15.
- Froyd & Lambert (1989).

- Hay, W.M., Hay, L.R., Angle, H.V. & Nelson, R.O. (1979). The reliability of problem identification in the behavioural interview. *Behavioural Assessment*, 1, 107–118.
- Hay, L.R., Nelson, R.O. & Hay, W.M. (1980). Methodological problems in the use of participant observers. *Journal of Applied Behaviour Analysis*, 13, 501–504.
- Iacono, W.G. (1991). Psychophysiological assessment of psychopathology. *Psychological Assessment*, 3, 309–320.
- Jacobson, N.S. (1992). Behavioural couple therapy: A new beginning. *Behaviour Therapy*, 23, 493–506.
- Lang, P.J. (1971). The application of psychophysiological methods to the study of psychotherapy and behaviour modification. In Bergin, A.E. & Garfield, S.L. (Eds.), *Handbook of Psychotherapy and Behaviour Change*. New York: Wiley.
- Lipinski, D.P. & Nelson, R.O. (1974). The reactivity and unreliability of self-recording. *Journal of Consulting and Clinical Psychology*, 42, 118–123.
- Kazdin, A.E. (1974). Self-monitoring and behaviour change. In Mahoney, M.J. & Thorsen, C.E. (Eds.), *Self-control: Power to the Person* (pp. 218–246). Monterey California: Brooks-Cole.
- Kazdin, A.E. (1979). Unobtrusive measures in behavioural assessment. *Journal of Applied Behaviour Analysis*, 12, 713–724.
- Korotitsch, W.J. & Nelson-Gray, R.O. (1999). Self-report and physiological measures. In Hayes, S.C., Barlow, D.H. & Nelson-Gray, R.O. *The Scientist-Practitioner: Research and Accountability in the Age of Managed Care*. (2nd ed., pp. 320–352). New York: Allyn & Bacon.
- Matarazzo, J.D. (1983). The reliability of psychiatric and psychological diagnosis. *Clinical Psychology Review*, 3, 103–145.
- Taylor, S. (1999). Behavioural assessment: review and prospect. *Behaviour Research and Therapy*, 42, 118–123.
- Watkins, C.E., Campbell, V.L., Nieberding, R. & Hallmark, R. (1995). Contemporary practice of psychological assessment by clinical psychologists. *Professional Psychology: Research and Practice*, 26, 54–60.
- Weinrott, M. & Jones, R.R. (1984). Overt versus covert assessment of observer reliability. *Child Development*, 55, 1125–1137.

William J. Korotitsch and
Rosemary O. Nelson-Gray

Related Entries

THEORETICAL PERSPECTIVE: BEHAVIOURAL, THEORETICAL PERSPECTIVE: COGNITIVE-BEHAVIOURAL, OBSERVATIONAL METHODS (GENERAL), OBSERVATIONAL TECHNIQUES IN CLINICAL SETTINGS, SELF-REPORTS (GENERAL), SELF-REPORTS IN BEHAVIOURAL CLINICAL SETTINGS, PSYCHOPHYSIOLOGICAL EQUIPMENT & MEASUREMENTS.

B

BEHAVIOURAL SETTINGS AND BEHAVIOUR MAPPING

INTRODUCTION

At first it might seem that behaviour settings and behavioural mapping are two separate and unrelated methods. Yet the true meaning of behaviour setting is that all behaviour is linked to a particular time and place; so any behavioural map is simply a record of behaviour that has always to be within a behaviour setting. In a very literal sense behavioural mapping is really the footprint of a behaviour setting or settings.

For those unfamiliar with the term, 'behaviour setting,' it refers to a standing pattern of behaviour which is tied to a particular place and time, (these) are simply the easily observed events of everyday life like the grocery store, the lawyer's office, 3rd grade class. They can be observed to begin at a regular time and end at a regular time and contain a recognized pattern of behaviour which is constantly repeated. If it is unclear whether settings which are adjacent in time or place are really separate, the K-21 scale is used. This scale is available in Barker and Wright (1955), Schoggen (1989) or Bechtel (1997). The central idea is overlap of population and behaviour. If there is more than a fifty per cent overlap on the seven scales (population, space used, leadership, objects, action, time, mechanisms) the putative settings are really one. The score of 21 is arbitrarily chosen as the cut off point to separate two units but any score between 17 and 23 can indicate some boundary problems (Bechtel, 1977). of observed human behaviour. They are the units into which humans sort themselves to get the daily business of living done.

Behavioural mapping is the narrower recording of specific behaviours within settings. A behavioural map (Ittelson, Proshansky & Rivlin, 1976) is a recording of where behaviour takes place on a floor plan of the setting, providing a two-dimensional record of the behaviour.

In special cases it is also possible to record the behaviour automatically (Bechtel, 1967). Behavioural maps can include more than one behaviour setting.

BEHAVIOUR SETTINGS AS ASSESSMENT TOOLS

A behaviour setting census, that is, a complete count of behaviour settings in a community over a year, is used to assess either a community or an individual. Community assessment is done by counting the number of behaviour settings (with their population numbers) that occur in a defined community for one year. Assessment of an individual is done by collecting the *behavioural range*, the number of settings an individual enters in a year or a shorter time span, depending on the purpose of the assessment. A year is necessary in order to include the kinds of settings which only occur once a year like Christmas Eve, Easter, Fourth of July, etc. Merely counting the number of settings can provide a measure of health for both communities and persons.

A healthy community can be defined as one that provides an adequate, or, preferably, more than adequate number of resources for its inhabitants. Healthy communities have about two settings available for each inhabitant. But there are other aspects which can be deduced from these numbers. For example, when two communities were compared (Barker & Schoggen, 1973), it was observed that one, a midwest town, had more behaviour settings available per child than a town in Great Britain. This was explained by the different philosophies on child rearing that existed in the two communities. In the midwest town it was assumed that the best way to rear children was to get them participating in adult life as soon as they could even though they might not be capable of performing at the adult level. In the

British community children were withheld from participation until it was deemed they were capable of participating at a reasonably competent level. The result was the midwest town had twice as many settings where children were present. If one agreed that the midwest philosophy was more valid, then the greater participation would be a measure of a healthy environment for children and (could even quantify the number of settings available to children vs. number of children and be used to evaluate goals.) Organizations can be assessed by use of behaviour settings. For example, in the study of school size (Barker & Gump, 1964), it was discovered that large schools had twenty times as many students as small schools but only five times as many settings. The consequence of this was that small schools can have twice the participation level of large schools in extra curricular activities, simply because there is more activity per student. The psychological consequences of this size discrepancy are also critical. Small schools report more satisfaction, competence, being challenged, engaging in important actions (leadership), being involved, achieving more cultural and more moral values. By contrast, large schools report more vicarious enjoyment (passive roles), large affiliation, and learning more about the school and persons in it.

Leadership is another important variable that can be measured when taking a behaviour-setting census. A simple scale is used for each behaviour setting; six is applied to a leader without whom the setting could not take place. A one person radio station is an example. A teacher of some strange language like Urdu might be another. There are few truly six-rated settings because most are shared leadership. For example, any organization with a vice president has a shared leadership. Most settings have leaders rated at the five level. Fours are officials like secretaries, treasurers, board members, etc. Anyone who has a role above that of plain member is a four.

Even a janitor is a four. Threes are the bona fide members of the setting who are not officials in any way. Twos are visitors to the setting and ones are onlookers outside the setting looking in. Sidewalk superintendents are ones. This simple leadership breakdown for every setting can be used to calculate the leadership roles available per person in a town. To return to our midwest versus British towns, the midwesterners had control of four times

as many settings as the British. This was because the British town had many outside persons entering and controlling settings. It is obvious that this simple scale can also be used as a measure of opportunity. For example, in small versus large military bases in Alaska (Bechtel, 1977) it was shown that had a much better chance at leadership roles in the smaller bases.

Using the Behavioural Range can be an effective way to measure the involvement of a patient outside of therapy (Bechtel, 1984). This method can be used to measure an average day, a month or a whole year in a person's life. The entire year is an accurate measure of lifestyle. For a quicker assessment, the therapist merely asks about the number of activities the patient engages in and assumes these are settings. Also critical is the role the patient takes in each setting, whether passive or some form of leadership. Two alcoholic women, aged 50, were assessed in this manner. The first patient had very few activities and when she saw how sparse her day was, remarked, 'Gee, I don't do very much.' Part of the therapy contract was to get her involved in more activities. The second patient made a rather impressive list of activities and was surprised by the breadth of engagement. Her contract was to use these settings as a better resource. The fact of participation was itself reassuring, however.

The Behavioural Range can also be used as a personal leadership measure by using the 1–6 scale for participation in each behaviour setting.

BEHAVIOUR MAPPING AS AN ASSESSMENT TOOL

Behavioural mapping was first used by Ittelson (1961) in a mental ward of a veteran's hospital (see Table 1). Patients in the ward were directly observed and their movements and behaviour coded on a floor plan of the ward. Cherulnik (1993) provides two examples of behavioural mapping used to evaluate changes in mental hospital wards. In both cases physical arrangements were modified to allow patients a closer proximity in order to encourage social interaction. And in both cases this was successful because pre–post behavioural mapping showed significant increases in social interaction.

A scale drawing of the place to be measured is first necessary with each physical feature labelled.

Table 1. Behavioural mapping categories from a mental ward (from Ittelson et al., 1976: 344)

Behaviour	Observational categories	Analytic categories
Patient reclines on bench, hand over face, but not asleep		
Patient lies in bed awake	Lie awake	
Patient sleeps on easy chair		
One patient sleeps while others are lined up for lunch	Sleeping	Isolated passive
Patient sits smiling to self	Sitting alone	
Patient sits, smoking and spitting		
Patient writes a letter on bench		
Patient takes notes from book	Write	
Patient sets own hair	Personal hygiene	
Patient sits, waiting to get into shower		
Patient reads newspaper and paces		
Patient reads a book	Read	Isolated active
Patient and nurse's aid stand next to alcove		
Patient stands in doorway smoking	Stand	
Patient paces between room and corridor		
Patient paces from room to room saying hello to other patients	Pacing	
Upon receiving lunch some patients take it to bedroom		
Patient sits at table and eats by self	Eating	
Patient cleans the table with sponge		
Patient makes bed	Housekeeping	
Two patients listen to record player	Phonograph-Radio	Mixed active
Patient turns down volume on radio		
Patient knits, sitting down	Arts	
Patient paints (oils), sitting down	And crafts	
Patient and registered nurses watch TV, together		
Patient watches TV, goes to get towel, returns	TV	
Patient stands and watches card games	Watching an activity	
Patient sits on cans in hall watching people go by		
Patient play soccer in corridor		
Patient and doctors play chess	Games	
One patient talks to another in reassurance tones		Social
Four patients sit facing corridor, talk sporadically		
Patient fails to respond to doctor's questions	Talk	
Patient introduces visitors to other patient		
Patient stands near room with visitors	Talk (visitor)	Visit
Patient comes in to flick cigrattes ashes		
Patients go to solarium	Traffic	Traffic

The categories of behaviour should be observable and codable. One problem is the intrusiveness of observers. Usually observers are introduced as 'architecture students' who want to observe how the design elements are used. Another problem is the time sampling. If architectural features are to be evaluated, the time of maximum use must first be determined. Another aspect, however, may be the span of time where the features being studied are not used at all. Use and disuse are often

problems with the same design feature. For example, in one study of a hospice done by Bill Ittelson and I, it was discovered that the chapel of the hospice was seldom used compared to other places. But when patients and staff were quizzed, it became apparent that the symbolic importance of the chapel made it more important than the actual use. An advantage of both the behaviour setting and behavioural mapping techniques is that they are essentially atheoretical and can be

used to test any theory that proposes to influence behaviour or design.

FUTURE PERSPECTIVES AND CONCLUSIONS

The use of behaviour settings and behavioural mapping continue in many post-occupancy and other evaluation studies. It is often the practice to include them as part of several methods in post-occupancy evaluation (POEs). However, many of the quantitative scales are often not used because researchers are not aware of their utility. For example, in several studies, the K-21 scale was used to measure boundary problems between two settings located adjacent to each other (see Barker, 1968 and Bechtel, 1987). Many times this scale can answer the question of whether a wall should be constructed between the settings.

The future of these measures is potentially greater than ever. The kind of data obtained is more readily understood by architects and engineers because it measures easily observed phenomenon (settings) which any layman can see and relate to. I can remember a conversation with Burgess Ledbetter, one of the architects I have worked extensively with in past years. He was designing a church of 10,000 members. I asked how he went about such an enormous task. He replied without hesitation, 'I just counted the potential behaviour settings.'

References

- Barker, R. & Gump, P. (1964). *Big School, Small School*. Palo Alto, CA: Stanford University Press.
- Barker, R. & Wright, H. (1955). *Midwest and its Children*. Evanston, IL: Row, Peterson.
- Barker, R. & Schoggen, P. (1973). *Qualities of Community Life*. San Francisco: Jossey Bass.
- Bechtel, R. (1967). Hodometer research in museums. *Museum News*, 45, 23–26.
- Bechtel, R. (1977). *Enclosing Behaviour*. Stroudsburg, PA: Dowden, Hutchinson & Ross.
- Bechtel, R. (1984). Patient and community, the ecological bond. In O'Connor, W. & Lubin, B. (Eds.), *Ecological Approaches to Clinical and Community Psychology*. 216–231. New York: Wiley
- Bechtel, R. (1997). *Environment & Behaviour, An Introduction*. Thousand Oaks, CA: SAGE Publications.
- Cherulnik, P. (1993). *Applications of Environment-Behaviour Research*. Cambridge: Cambridge University Press.
- Ittelson, W. (1961). *Some Factors Influencing the Design and Function of Psychiatric Facilities*. (Progress Report) Brooklyn, NY: Brooklyn College.
- Schoggen, P. (1989). *Behaviour Settings*. Stanford: Stanford University Press.

Robert B. Bechtel

Related Entries

BEHAVIOURAL ASSESSMENT TECHNIQUES, OBSERVATIONAL METHODS (GENERAL), PERSON/SITUATION (ENVIRONMENT) INTERACTION, THEORETICAL PERSPECTIVE: BEHAVIOURAL

B | BIG FIVE MODEL ASSESSMENT

INTRODUCTION

The Big Five model of personality traits derives its strength from two lines of research, the psycholexical and the factoranalytic tradition, from which the interchangeably used names Big Five model and Five Factor Model respectively originate. The two traditions have produced remarkably similar five-factor structures that mark a point of no return for personality

psychology. An extensive review of history and theory with respect to the Big Five can be found in De Raad (2000).

The Big Five factors have been endorsed with a distinctive status, derived from the extensive, omnibus-character of the underlying *psycholexical* approach, and based on two characteristics, namely its *exhaustiveness* in capturing the semantics of personality and its recourse to *ordinary language*. Though both these

characteristics may be improved upon, in comparison to other approaches to personality, the psycholexical approach outranks semantic coverage, and it has optimized the level of communication on personality traits by faring merely on readily intelligible units of description.

The model has served as a basis for the development of assessment instruments of various kinds. In the following paragraphs, different assessment forms based on the Big Five model, as well as some representative assessment systems are briefly described, including Big Five trait-markers, Big Five inventories, and some instruments that have been moulded after the Big Five framework. To begin with, a brief content description of the Big Five constructs is given.

THE BIG FIVE CONSTRUCTS

The Big Five constructs, Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Intellect/Autonomy, made a long journey, covering about a whole century, towards a strong performance in the psychological arena during the last decade of the twentieth century. A straight count of the references made to each of the presently identified Big Five constructs in abstracts since 1887, tells that of the total number of 17,262 references made, Extraversion (and Introversion) and Neuroticism (and Emotional Stability) are the absolute winners, with 8574 and 6189 references respectively. This picture sustains the historical 'Big Two' of temperament (Wiggins, 1968). The historical third, Intellect, with 1534 references, may refer to both traits and abilities.

Extraversion and Introversion

No single pair of traits of personality has been quite so widely discussed and studied as that of Extraversion and Introversion. Their main understanding at the onset of their appearance was Jungian. To Jung Extraversion is the outward turning of psychic energy toward the external world, while Introversion refers to the inward flow of psychic energy towards the depths of the psyche. Extraversion is denoted by habitual outgoingness, venturing forth with careless confidence into the unknown, and being particularly interested in people and events in the external world.

Introversion is reflected by a keen interest in one's own psyche, and often preferring to be alone.

Extraversion is a dimension in almost all personality inventories of a multidimensional nature, which in fact sustains its relevance and its substantive character. Moreover, many studies have provided behavioural correlates of this construct, such as the number of leadership roles assumed, and frequency of partying. Extraversion has also been found relevant in contexts of learning and education (De Raad & Schouwenburg, 1996) and of health (e.g. Scheier & Carver, 1987).

Agreeableness

Agreeableness is the personality dimension with the briefest history. Yet, while longtime constructs as Love and Hate, Solidarity, Conflict, Cooperation, Kindness, which are part and parcel of this dimension, may have been pivotal to the organization of social life throughout the history of mankind, as a personality dimension it essentially popped up with the rise of the Big Five. Agreeableness can be considered as being dominated by 'communion', the condition of being part of a spiritual or social community. Graziano and colleagues have described the details of the history of this construct (e.g. Graziano & Eisenberg, 1997).

Agreeableness is argued to play a role as a predictor of training proficiency (e.g. Salgado, 1997). In health psychological research, Agreeableness plays a documented role. Coronary heart disease is more likely to develop in competitive and hostile people than in those who are more easygoing and patient (cf. Graziano & Eisenberg, 1997).

Conscientiousness

Conscientiousness has been drawn upon as a resource in situations where achievement is of important value, that is, in contexts of work, learning and education. The construct represents the drive to accomplish something, and it contains the characteristics necessary in such a pursuit: being organized, systematic, efficient, practical, and steady.

Conscientiousness is found to be consistently related to school performance (e.g. Wolfe & Johnson, 1995), and job performance (e.g. Salgado, 1997).

Emotional Stability and Neuroticism

The first inventory measuring neurotic tendencies is Woodworth's (1917) Personal Data Sheet, developed to assess the ability of soldiers to cope with military stresses. Thurstone and Thurstone (1930) developed a neurotic inventory called 'A Personality Schedule' to assess the neurotic tendencies of university freshmen. As one of the 'Big Two', Neuroticism (or 'Anxiety') had been observed by Wiggins (1968) most notably in several of the works of Eysenck, Cattell, Guilford, and Gough.

Neuroticism has been found relevant as a predictor of school attainment (e.g. Entwistle & Cunningham, 1968). In the clinical situation, neuroticism is found relevant in the assessment of personality disorders (cf. Schroeder, Wormworth, & Livesley, 1992). Neuroticism correlates significantly with measures of illness (e.g. Larsen, 1992).

Intellect and Openness to Experience

Feelings are usually running highest for the Fifth of the Big Five. This refers to its naming but also to its origin and its relevance as a personality trait factor. Discussions with respect to this factor incorporate the various points of criticism that are expressed over the Big Five as a model. Several candidates for factor five have been suggested, such as Culture, Intellectance, and Openness to Experience (see De Raad, 2000).

In assessment situations the Fifth of the Big Five may be relevant in psychiatry and clinical psychology. Aspects of Openness to Experience seem to be related to several disorders (Costa & Widiger, 1994). In contexts of learning and education, Openness to Experience has been related to learning strategies. Learning strategies possibly mediate a relationship between Openness to Experience and grade point average (cf. Blickle, 1996).

FACETS OF THE BIG FIVE

The Big Five factors represent an abstract level of personality description that may capture specificity at a lower level. Perugini (1999) distinguishes

two ways to specify different levels of abstractness, a hierarchical and a circumplex approach. The *hierarchical* approach considers facets as first order factors and the Big Five as second order factors. The *circumplex* approach represents a fine-grained configuration in which facets are constituted as blends of two factors, based on the observation that many traits are most adequately described by two (out of five) substantial loadings. Because of its explicit coverage of the trait domain, the latter model provides an excellent starting point for the development of personality assessment instruments.

BIG FIVE TRAIT-MARKERS

Possibly the most direct way to arrive at an instrument assessing the Big Five is to select trait-variables as *markers* of the Big Five, on the basis of their loadings on those factors. Simply taking the first *n* highest loading trait-variables per factor might do the job. A frequently used marker list to measure the Big Five is the one described in Norman (1963). The list is based on earlier work by Cattell (1947). For the history of this and similar constructs from the same period, as well as for a comprehensive coverage of many psycholexical studies, see De Raad (2000). Goldberg (1992) developed an adequate list of 100 'unipolar' markers for the Big Five. In his 1992 article Goldberg concludes: 'It is to be hoped that the availability of this easily administered set of factor markers will now encourage investigators of diverse theoretical viewpoints to communicate in a common psychometric tongue'.

BIG FIVE INVENTORIES AND QUESTIONNAIRES

Several instruments have been developed to assess the Big Five factors. Besides those that are briefly described in the following sections, a few others should be mentioned such as the BFI (John, Donahue, & Kentle, 1991), the HPI (Hogan & Hogan, 1992), the IPIP (Goldberg, 1999) and the HiPIC (Mervielde & De Fruyt, 1997). A few characteristics of some main Big Five instruments are summarized in Table 1.

Table 1. Summary of the some main big five inventories

Instrument	Authors	Factors					Variables
		I	II	III	IV	V	
100 Unipolar markers	Goldberg	Extraversion/ surgency	Agreeableness	Conscientiousness	Emotional stability	Intellect	100 adjectives
FFPI	Hendriks, Hofstee, De Raad	Extraversion	Agreeableness	Conscientiousness	Emotional stability	Autonomy	100 items
NEO-PI-R	Costa, McCrae	Extraversion	Agreeableness	Conscientiousness	Neuroticism	Openness to experience	240 items
BFQ	Caprara, Barbaranelli, Borgogni, Perugini	Energy	Friendliness	Conscientiousness	Neuroticism	Openness	120 + 12 Lie items
FF-NPQ	Paunonen, Ashton, Jackson	Extraversion	Agreeableness	Conscientiousness	Neuroticism	Openness to experience	60 nonverbal items

FFPI (Five Factor Personality Inventory)

This inventory (Hendriks, Hofstee, & De Raad, 1999) is unique in several respects. It took its starting point in the circumplex approach with the so-called Abridged Big Five Circumplex (AB5C), distinguishing 90 facets that provide an optimal coverage of the semantics of the Big Five system. The pool of 914 items that was agreed upon to represent the AB5C system, was made available with approximately identical phrasings in Dutch, German, and English. Items were only accepted for the final pool if clear, unambiguous translations in those languages could be found. The final instrument, comprising 100 items, 20 for each of the five scales, is trilingual in nature. The items have a simple and easy to understand behavioural format, put in third person singular, which makes them suitable for both other-ratings and self-ratings. Some examples of items are: Has a good word for everyone, Makes friends easily, Suspects hidden motives in others, Makes people feel uncomfortable, Feels at ease with people, Shows his/her feelings, Gives compliments, and Respects others. Besides scores for the Big Five dimensions, the FFPI enables the computation of an additional 40 bipolar facet scores, derived as blends of the Big Five.

NEO-PI-R (NEO Personality Inventory Revised)

Costa and McCrae's (1992) NEO-PI-R is the most frequently used personality questionnaire to assess the Big Five. The development of the N (Neuroticism), E (Extraversion), and O (Openness to Experience) scales started with a cluster analyses of 16PF data, yielding two clusters called 'Adjustment-Anxiety' and 'Introversion-Extraversion', and a third cluster conceptualized as an Experiential Style dimension (openness versus closedness to experience). After taking knowledge of an early Big Five formulation, Costa and McCrae added Agreeableness and Conscientiousness to their three-dimensional system, assuming that their three dimensions, including Openness to Experience, captured the first three of the Big Five. Costa and McCrae's first Big Five version (the NEO-PI) included scales to assess six facets of Neuroticism, Extraversion, and Openness to Experience. Only the 240-item

NEO-PI-R (Costa & McCrae, 1992) also included six facets of Agreeableness and Conscientiousness.

BFQ (Big Five Questionnaire)

The Big Five Questionnaire (BFQ; Caprara, Barbaranelli, Borgogni, & Perugini, 1993) has been developed using a top down approach, by first defining the five dimensions, and subsequently defining the most important facets for each dimension. The BFQ was developed alongside the first psycholexical study in Italian and some of its findings were taken into account, especially, to define the first factor. Accordingly, in the BFQ, the first factor is defined as Energy – rather than as Extraversion. The BFQ is easily administered and includes unique features such as a relatively small number of items (120) and scales to assess two facets per factor; in addition, it provides a Social Desirability response set scale of 12 items. Recently, a children version (BFQ-C, 65 items) has also been developed.

FF-NPQ (Five-Factor Nonverbal Personality Questionnaire)

A controversy with respect to verbal self- and other-ratings is that they may reflect consistencies in language rather than consistencies in observed behaviour. For this reason, Paunonen, Ashton, and Jackson (2001) developed an instrument that did not make use of verbal items, but included cartoon-like pictures, in which a person performs specific behaviours in specific situations. The investigators initially developed a nonverbal item pool for a person perception study and aiming to represent traits of Murray's system of needs. From this item pool a subset of 136 items was selected to form the Nonverbal Personality Questionnaire (NPQ) measuring 16 personality traits. With a few exceptions items were selected from the NPQ to form the 60-item FF-NPQ, with 12 items measuring each of the Big Five factors. This instrument takes about 10 minutes to finish.

QUESTIONNAIRES RELATED TO OR MOULDED AFTER THE BIG FIVE

The impact of the Big Five factors have been such that researchers often clarify the relations of their

own alternative trait models with the Big Five. A few such alternative models have been proposed, such as a Big Three (Peabody & Goldberg, 1989), a Big Six (Jackson, Ashton, & Tome, 1996), a Big Seven factor model (Almagor, Tellegen, & Waller, 1995) and an alternative Five Factor model (Zuckerman, 1994). All these models share features with the Big Five but differ too.

In addition, some classic instruments to assess important personality dimensions have been moulded after the Big Five. Typically, this implied the development of a new coding format for existing items in those instruments so as to yield a measure of the Big Five factors. Examples of such instruments are the ACL (FormyDuval et al., 1995) and the 16PF (Hofer, Horn, & Heber, 1997). A more specific situation is provided by the recoding of the MMPI-2 into the *Personality Psychopathology-Five Questionnaire* (PSY-5). The MMPI is one of the most used personality inventories for psychopathological assessment, originally developed in the 1940s and recently refurbished (MMPI-2). Harkness and McNulty (1994) developed the so-called PSY-5 constructs starting from a pool of symptoms and characteristics of both normal and dysfunctional personality functioning leading to the identification of 60 major topics in human personality. These topics were used to generate five higher order aggregates that have some resemblance with the Big Five, with especially the fifth factor remaining evidently uncovered.

FUTURE PERSPECTIVES

Because the Big Five model has acquired the status of a reference-model, its uses can be expanded to that of systems of classification and clarification for descriptive vocabularies that are not developed from a Big Five perspective, in order to evaluate the comprehensiveness of the trait-semantic of those vocabularies. Examples of such uses are given in De Raad (2000). Moreover, the model is expected to play an important role in modern theory building, because its five main constructs capture so much of the subject matter of personality psychology. An example is Digman (1997), who succeeded in relating the Big Five factors to

a higher order schema which brings together central concepts from various theories from the history of personality psychology.

Many more instruments along the main Big Five theme will be developed in the near future, as translations of existing instruments or as instruments that are completely developed within particular languages. Especially efforts may be expected to specify facets of the Big Five that can be cross-culturally validated.

CONCLUSION

Trait structures from different languages differ, and so do assessment instruments, imported or not. This conclusion is not dramatic; it is a challenge to cross-cultural research-programmes to isolate and identify what is valid across cultural borders, and to specify the particulars of the different cultures. A lot has yet to be done. The Big Five factor model has shown to be highly prolific in the construction of assessment instruments, notwithstanding the fact that its significance has only been recognized during the last decade of the twentieth century. Moreover, the Big Five factors are far from definitive, and the derived assessment instruments deserve constant attention and an open eye for new facets and features to be included, in the model as well as in its assessment.

References

- Almagor, M., Tellegen, A. & Waller, N.G. (1995). The Big Seven model: a cross-cultural replication and further exploration of the basic dimensions of natural language trait descriptors. *Journal of Personality and Social Psychology*, *69*, 300–307.
- Blickle, G. (1996). Personality traits, learning strategies, and performance. *European Journal of Personality*, *10*, 337–352.
- Caprara, G. V., Barbaranelli, C., Borgogni, L. & Perugini, M. (1993). The Big Five questionnaire: a new questionnaire to assess the Five Factor Model. *Personality and Individual Differences*, *15*, 281–288.
- Cattell, R.B. (1947). Confirmation and clarification of primary personality factors. *Psychometrika*, *12*, 197–220.
- Costa, P.T., Jr. & McCrae, R.R. (1992). *Revised NEO Personality Inventory (NEO PI-RTM) and NEO Five-Factor Inventory (NEO-FFI) Professional Manual*. Odessa, FL: Psychological Assessment Resources.

- Costa, P.T., Jr. & Widiger, T.A. (Eds.) (1994). *Personality Disorders and the Five-Factor Model of Personality*. Washington, DC: American Psychological Association.
- De Raad, B. (2000). *The Big Five Personality Factors: The Psycholexical Approach to Personality*. Goettingen: Hogrefe & Huber Publishers.
- De Raad, B. & Schouwenburg, H.C. (1996). Personality in learning and education: a review. *European Journal of Personality*, 10, 303–336.
- Digman, J.M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology*, 73, 1246–1256.
- Entwistle, N.J. & Cunningham, S. (1968). Neuroticism and school attainment – a linear relationship? *Journal of Educational Psychology*, 38, 123–132.
- FormyDuval, D.L., Williams, J.E., Patterson, D.J. & Fogle, E.E. (1995). A 'big five' scoring system for the item pool of the Adjective Check List. *Journal of Personality Assessment*, 65, 59–76.
- Goldberg, L.R. (1992). The development of markers of the Big-Five factor structure. *Psychological Assessment*, 4, 26–42.
- Goldberg, L.R. (1999). A broad-bandwidth, public-domain, personality inventory measuring the lower-level facets of several five-factor models. In Mervielde, I., Deary, I., De Fruyt, F. & Ostendorf, F. (Eds.), *Personality Psychology in Europe*, Vol. 7 (pp. 7–28). Tilburg, The Netherlands: Tilburg University Press.
- Graziano, W.G. & Eisenberg, N. (1997). Agreeableness: a dimension of personality. In: Hogan, R., Johnson, J. & Briggs, S. (Eds.), *Handbook of Personality Psychology* (pp. 795–824). San Diego, CA: Academic Press.
- Harkness, A.R., McNulty, J.L. & Ben-Porath, Y.S. (1995). The Personality Psychopathology Five (PSY-5): constructs and MMPI-2 scales. *Psychological Assessment*, 7, 104–114.
- Hendriks, A.A.J., Hofstee, W.K.B. & De Raad, B. (1999). The Five-Factor Personality Inventory (FFPI). *Personality and Individual Differences*, 27, 307–325.
- Hofer, S.M., Horn, J.L. & Eber, H.W. (1997). A robust five-factor structure of the 16PF: strong evidence from independent rotation and confirmatory factorial invariance procedures. *Personality and Individual Differences*, 23, 247–269.
- Hogan, R. & Hogan, J. (1992). *Hogan Personality Inventory Manual*. Tulsa, OK: Hogan Assessment Systems.
- Jackson, D.N., Ashton, M.C. & Tomes, J.L. (1996). The six-factor model of personality: facets from the Big Five. *Personality and Individual Differences*, 21, 391–402.
- John, O.P., Donahue, E.M. & Kentle, R.L. (1991). *The Big Five Inventory – Version 4a and 54*. Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research.
- Larsen, R.J. (1992). Neuroticism and selective encoding and recall of symptoms: evidence from a combined concurrent–retrospective study. *Journal of Personality and Social Psychology*, 62, 480–488.
- Mervielde, I. & De Fruyt, F. (1997). *Hierarchical Personality Inventory for Children (HiPIC) aged 6 to 12*. Lisse: Swets and Zeitlinger.
- Norman, W.T. (1963). Toward an adequate taxonomy of personality attributes: replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology*, 66, 574–583.
- Paunonen, S.V., Ashton, M.C. and Jackson, D.N. (2001). Nonverbal assessment of the Big Five personality factors. *European Journal of Personality*, xx, xxx–xxx.
- Peabody, D. & Goldberg, L.R. (1989). Some determinants of factor structures from personality trait-descriptors. *Journal of Personality and Social Psychology*, 57, 552–567.
- Perugini, M. (1999). A proposal for integrating hierarchical and circumplex modelling in personality. In Mervielde, I., Deary, I., De Fruyt, F. & Ostendorf, F. (Eds.), *Personality Psychology in Europe*, Vol. 7 (pp. 85–99). Tilburg, The Netherlands: Tilburg University Press.
- Salgado, J.F. (1997). The five factor model of personality and job performance in the European Community. *Journal of Applied Psychology*, 82, 30–43.
- Scheier, M.F. & Carver, C.S. (1987). Dispositional optimism and physical well-being: the influence of generalized outcome expectancies on health. *Journal of Personality*, 55, 169–210.
- Schroeder, M.L., Wormworth, J.A. & Livesley, W.J. (1992). Dimensions of personality disorder and their relationships to the Big Five dimensions of personality. *Psychological Assessment*, 4, 47–53.
- Thurstone, L.L. & Thurstone, T.G. (1930). A neurotic inventory. *Journal of Social Psychology*, 1, 3–30.
- Wiggins, J.S. (1968). Personality structure. *Annual Review of Psychology*, 19, 293–350.
- Woodworth, R.S. (1917). *Personal Data Sheet*. Chicago: Stoelting.
- Wolfe, R.N. & Johnson, S.D. (1995). Personality as a predictor of college performance. *Educational and Psychological Measurement*, 55, 177–185.
- Zuckerman, M. (1994). An alternative five factor model of personality. In Halverson, C.F., Kohnstamm, G.A. & Martin, R.P. (Eds.), *The Developing Structure of Temperament and Personality from Infancy to Adulthood* (pp. 53–68). New York: Erlbaum.

Boele de Raad and Marco Perugini

Related Entries

PERSONALITY (GENERAL)

BRAIN ACTIVITY MEASUREMENT

INTRODUCTION

Electroencephalograms (EEGs) from the human scalp were first recorded in 1924 by Hans Berger. It is assumed that they are generated by brain activity related to information processing. EEG is mainly caused by nerve cell activity, whereas other brain imaging methods are more related to blood flow and metabolic parameters. Moreover, the direct coupling of EEG with biological flow of information allows a continuous and chronometric approach to the basis of cognitive processing. Variations of EEG require synchronous and massive parallel activity in wide-ranging populations of neurons and the measures are done in a great distance to the generators. Thus spatial resolution is less than in other brain imaging techniques.

Actually EEG potentials occur in several locations with alternating polarity. This finding is consistent with models of information processing assuming separate modules of cognitive functioning, which interact continuously in terms of uptake, processing and passing on of information.

The main fields of the psychological use of EEG are in *cognition*, in search of cognitive relevant modules in the brain and their temporal interaction. Distortion of common spatial or temporal regularity in potential dynamics (such as dimensional complexity) can be interpreted as a sign of uncommon or *emotional processing*. Brain activity is present when awake as well as during sleep, in which a number of *sleep stages* and sleep parameters can be differentiated by using certain criteria. Deviant patterns of EEG activity can be used to characterize *psychopathological states* or could be caused by *drug effects*.

PARAMETERS

Neurophysiological Basis

It is widely accepted that most of the time both, excitatory and inhibitory postsynaptic potentials

simultaneously are present in the pyramidal cells of the upper and middle cortical layers. Usually they are in balance without releasing considerable action potentials. It is assumed that this is particularly true when a module became charged without immediate output. A negative potential on the surface is measured because excitatory synapses are predominant in upper layers (negative interstitium in the upper layers). The release of action potentials (negative interstitium far below) will change the dipole causing a positive potential.

Basic Activity

Negative and positive potentials in EEG alternate with main fluctuations within about 0.1 s (equivalent around 10 Hz). Dominant frequencies in the range of about 8–12 Hz are called *EEG alpha*. Alpha is observed in awake but resting subjects without demanding memory load. Alpha is generated by burst activity produced by loops between thalamic nuclei and the related cortical areas in case of attenuated stimulation. A lower portion of the alpha band (8–10 Hz) is discussed as reflecting attenuation of cortical activity during mental load while attending stimuli actively, for example in a time series resulting in partial loss of feature-related activation. The upper portion of alpha seems to be closer related to a more general attenuation of mental load mainly in processing stimuli, even by exogenous stimulation. Frequencies of 12–14 Hz (*EEG spindles*) seem to be indicative of active suppression of sensory stimuli during sleep.

Frequency 4–8 Hz (*EEG theta*) is discussed as indicative for extension of receptive fields, for example in coarse classification of stimuli. Theta is found to be increased during drowsiness and undirected memory search (flight of thought) as well as during top down or effortful processes causing directed memory search. The latter findings gave rise to the view that theta reflects

involvement of hippocampal memory functions. Theta power can be found in posterior locations as well as above the premotor cortex indicating activated wide motor concepts. In learning response concepts, frontal theta is increased in good learners compared to poor learners.

Frequencies < 4 Hz (*EEG delta*) are found in slow wave sleep. Frequencies in the range of 40 Hz (*EEG gamma*) correspond to activities of neuronal ensembles, where some particular stimulus features are bound together building up a cognitive representation of an object or a gestalt. It is discussed that frequencies of 6 Hz may give rise to about seven oscillations of 40 Hz representing about seven distinct information chunks per second. There is a broad range of irregular frequencies between 14 and 40 Hz contributing to the shape of raw EEGs of awake subjects, which is called *EEG beta*.

Analysis of EEG basic activity needs data processing in the frequency domain and is useful for characterizing widespread cortical processes. It can be done for any time range as conceded by resolution and lower limit of the frequencies of interest, such as for mental states or for epochs chosen in relation to certain events.

Event-Related Potentials

Information processing can often be related to external events, such as the onset of a stimulus or a response. EEG potentials in the time domain corresponding to assumptions on expecting or processing of stimuli as well as preparing or evaluating of responses allow a kind of mental chronometry.

The most common potential observed before stimulus onset is a *contingent negative variation* (CNV) in the case of so-called imperative stimuli (which request fast responses) revealing increasing motor preparation. Consecutive to the onset of a stimulus, negative potentials reflect the load of certain brain areas stimulated by individual significant stimulus features. Physical features produce a load in modal specific areas in a time range of about 150 ms after onset, called *processing negativity* (N1). More abstract or related features lead to a load mainly depending on context information, for example in case of similar stimuli, in the context of a task or in case of other kinds of involuntary or voluntary attention. Under these circumstances mental

load mostly can be interpreted as a kind of mismatch and the related potential in the time range between 200–300 ms after onset is called *mismatch negativity* (N2).

Information load in individual brain areas is mostly followed by passing forward information to related or higher order areas, as revealed by a positive potential. An early positive deflection P1 (circa 100 ms) reflects forward processing of prepared (biological or overlearned) stimulation. A positive potential P2 (circa 200 ms) in the time range between N1 and N2 could be interpreted as forward processing from physical to psychological relevant features. Extraction of the psychological content ('semantics') means classification and relating to an abstract concept. Forward processing after this by a P3 or P300 (> 300 ms after onset) is discussed as cognition of the stimulus in terms of upgrading of the hitherto model of the environmental context. While most of the processes up to P3, even automatic respondings, are unconscious, forward processing after mismatch is assumed to be obligatory for being aware of the stimulus.

Longer lasting processing increases P3 latency and widens the peak. Peak amplitude increases with task relevance and stimulus uncertainty. Important properties of stimulus processing can be studied by the odd ball paradigm. This paradigm consists of at least two classes of stimuli appearing randomly in time, where the instances of one stimulus are rare (20–30%) and a task should be done by using the rare stimuli (for example counting). Under these circumstances rare stimuli are responded by potentials with high P3s.

In case of mismatch of the extracted meaning of a stimulus compared to its semantic context, late negative and positive potentials can occur. *Semantic mismatch* occurs if a sentence ends with unexpected words or phrases (N400). Conducting information processing to a *reanalysis* is discussed in cases when a late positive potential follows (P600).

DATA ACQUISITION

The EEG Laboratory

EEG raw data have to be obtained in a laboratory protected against vibration and noise. Recording can be done without electric

field protection, if external field generators are weak and well known (50/60 Hz). Usually a separate space for subjects (including display and response devices) and acquisition apparatus (amplifier and monitoring) should be provided. The electrical potentials recorded from the scalp are of low amplitude and have to be preamplified close behind the electrodes and amplified by high quality amplifiers. A/D converter is used to convert the analogue time-continuous voltage-time series into a digitalized time-discrete signal. Analog-digital conversion rate (sampling rate) has to be at least twice as high as the highest frequency of interest in the signal to be measured to prevent the appearance of frequencies not present in the original signal. Preparation of derivation should be conducted by trained personnel. Otherwise all requirements, instruction, supervision and data acquisition by an examiner and/or computer has to be done as it is usual in psychological experiments.

Electrode-Skin Interface

In the brain, a great variety of processes takes place continuously continuously generating time varying (bio)electric potential fields over the scalp. EEG signals are voltage time series reflecting the potential difference between two field points derived from the scalp by electrodes. Analyses of human EEG are usually based on frequencies of 0 to 100 Hz containing magnitudes approximately of 0 to 200 μV .

Employing high input impedance EEG amplifiers, a variety of different electrode materials (Ag/AgCl, tin, silver, gold) in combination with electrode jelly may be used. Caps with embedded electrodes permit simple handling and replication. Impedances up to 40 k Ω are permitted (Ferree et al., 2001), but less than 5000 Ω are usually preferred. This can be attained by abrading slightly the surface or even scratching the skin surface with a sterile needle. However, injuries have to be avoided.

Points of Derivation

Referential recording is based on the assumption that one electrode site is an inactive reference site and the active site of interest is recorded with respect to that reference. Reference sites with

minor electrical activity such as the earlobes, mastoids, nose are preferred.

A reference-independent measure of the potential field is required for studying scalp topography and for source localization. One approach to overcome the reference-site problem is to use the so called average-reference using the mean of all recording channels at each time point to approximate an inactive reference. (Recording) problems arise because electrodes are not evenly distributed over the head surface. Another approach is to use reference-free transformations, such as current source density analysis (CSD) which is based on the second derivative of the interpolated potential distribution (Laplacian operator). The latter method accentuates local sources and masks interelectrode correlations. In order to get valid approximations, both approaches require a sufficient spatial electrode density.

Due to the prerequisites especially for successful topographic mapping and source localization a standardized system of electrode placement with up to 74 electrodes is usually used (10–10 system or ‘10% system’). Depending on certain research questions, a fewer number is used and/or interpolated sites are chosen. Advanced derivations use a 5% system with up to 345 electrodes (Oostenveld & Praamstra, 2001).

Common Steps in Artifact Rejection

The raw EEG signal may be contaminated by both technical (as power supply) and biological electric fields (electric activity of eyes, heart, muscle tension etc.). Parts of the EEG signal which are not generated by distinct brain processes are called ‘artifacts’. Artifacts are not easy detectable and there are no common methods of artifact rejection. Thus contaminations of the brain signals have to be avoided by careful planning of the derivation setting (avoiding technical carelessness and unnecessary muscular activity as well as eyeblinks). After derivation the experimenter should do some ‘eyeballing’ on the signals. With DC-derivations it is useless to define an amplitude criterion for rejection (for example +80 μV). Noisy parts of the signal should be removed. Within one experiment the same criteria have to be applied for all subjects. Correction of ocular artifacts

could be done in some cases (for uneasy children or patients) by use of special algorithms. Zero phase-shift low pass filters (about 20 Hz) are used for signal smoothing in ERP analyses.

DATA ANALYSIS

Signal Characteristics

Event-related potential (ERP) analysis is based on the assumption that part of the electrical brain activity is in a stable time relationship responding to a stimulus and the remaining brain activity is considered to be stationary noise. Hence segment-averaging is used to reduce variance depending on the ratio of time-locked to non-time-locked signal portion.

In general EEG signals are considered to be generated by stochastic processes with unknown probability density functions. Hence the processes are characterized by moments and moment functions. Usually EEG time series are studied up to second order of moments (mean, variance) and moment functions (covariance functions). *Higher-order statistics* (HOS) have to be used to analyse signal properties which deviate from Gaussian amplitude distribution (signal skewness, signal kurtosis). Given a signal of interest with non-zero HOS noised by Gaussian noise, then HOS is less affected by noise than second-order analyses.

Apart from the stochastic approach attempts have been made to describe EEG signals as the output of a complex *deterministic process* by use of nonlinear difference equations. The corresponding mathematical base originates from the field of 'deterministic chaos'. One frequently used measure is called *EEG dimensional complexity* (DCx) and this yields information regarding the complexity of processes in the brain.

Methods of Spectral Estimation

One widely used method when analysing EEG time series is spectral analysis, which means analysing a given signal with respect to its properties within the frequency domain. (Problems arise in analysing rapid amplitude changes within low frequency bands.)

Spectra can be obtained by filtering the signal with a set of narrow bandpass filters. This

procedure is common in determining *event-related desynchronization* (ERD) where activated cortical areas are assumed to be desynchronized compared to an idling state. After averaging over trials to discriminate between event-related and non-event-related power changes a standardized difference term between signal power in the analysed interval (A) and in a reference interval (R) is calculated:

$$\text{ERD} = ((R - A)/R) \times 100\%.$$

Fourier transform (FT) and wavelet transform (WT) are linear transformations of the signal from time to frequency domain. The most widely used approach for spectral estimation based on FT is the periodogram. Here the estimation is achieved by decomposing the signal recorded over time T in sines and cosines. To get reliable spectral estimates when analysing short epochs in the range of a few seconds (*short time Fourier transform* STFT), a correction of the data segments is required. This could be done by tapering functions in the time domain (for example Hanning window). Additionally segment-averaging or smoothing is used to reduce variance. Note that frequency resolution (in hertz) is inverse proportional to the epoch length T (in seconds). With STFT, dynamics over time can be displayed in a time-frequency plane.

The idea behind *wavelets* is simply to have more appropriate functions than sines and cosines when dealing with non-stationary impulse-like events (spikes and transients, for example high-frequency bursts and K-complexes). The principle way of wavelet analysis is to define a wavelet prototype function $W(t)$ as an analysis template. The corresponding wavelet basis, $W_{s,l}(t)$, is obtained from the mother wavelet $W(t)$ by varying the scaling parameter s and the locating parameter l . Thus, the wavelets $W_{s,l}(t)$ are time shifted (l) and scaled (s) derivations of $W(t)$. Each analysis template $W_{s,l}(t)$ represents a band pass function with a central frequency f_0 , localized in the time-frequency plane at $t = l$ and $f = f_0/s$. At any scale s the wavelet has not one frequency, but a band of frequencies, and the bandwidth is inverse proportional to s . The finer the resolution in time domain (small s) the less is the resolution in frequency domain and vice versa. The output can be displayed in the time-frequency plane

analogous to STFT, reaching a maximum when the signal of interest most resembles the analysis template. Summing up, it may be said that the short-time Fourier transform is well adapted for analysing all kinds of longer lasting oscillatory like waveforms, whereas wavelets are more suited for the analysis of short duration pulsations and for signal detection, for instance in ERPs.

With *model based methods of spectral estimation* the raw EEG is interpreted as the output of a linear filter excited by white noise. EEG signal modelling and hence spectral estimation is based on derivatives of the autoregressive moving average model (ARMA), which is described by a linear difference equation:

$$X_t = a_1 X_{t-1} + \dots + a_p X_{t-p} + b_1 e_{t-1} + \dots + b_q e_t - q + e_t \quad (1)$$

where p denotes autoregressive lags and q denotes moving average lags. Terms containing e characterize white noise.

A multitude of models similar to Equation (1) is used. All of them are based on assumptions concerning the underlying stochastic processes rather than describing a certain biophysical model. Successful spectrum estimation depends critically on the selection of the appropriate model, the model order and the fitting method for estimating the coefficients (for example least-square-methods, maximum-likelihood-methods). Model-based spectral estimation compared to the Fourier transform approach is useful when dealing with very short segments.

Generally Used Spectral Estimations

The *power spectrum* (auto-spectral density function) displays the signals distribution of variance or power over frequency. The *cross-power spectrum* (cross-spectral density function) reflects the covariance between two EEG channels as a function of frequency.

A frequently used quantity is the cross-power spectrum normalized by the autospectra, the so called *coherence spectrum* *Coh*. EEG coherence analysis is regarded as a tool for studying interrelationships with respect to power and phase between different cortical areas during a certain psychological manipulation (such as sen-

sory stimulation, voluntary movements). The values of the coherence function lie in the range from 0 to 1. It is assumed that a strong functional relationship between two brain regions is reflected by a high coherence value. To avoid trivial results (volume conduction) coherence should only be interpreted if the phase lag between the two channels is nonzero. Erroneous estimations may be caused for example by A/D converters producing artificial phase lag while sampling the data or by reference electrode effects.

The *bispectrum* B_i (the product of two spectra) and its normalized derivative *bicoherence* are third order measures in the frequency domain related to the signal skewness. They are tools for detecting the presence of non-linearity, particularly quadratic phase-coupling, i.e. two oscillatory processes generate a third component with a frequency equal to the sum (or difference) of two frequencies f_1 and f_2 . As compared to the power spectrum, more data is usually needed to get reliable estimates.

Non-Invasive Localization of Neuronal Generators

The EEG can be used as a method for functional neural imaging. Its advantage is to display dynamic brain processes on a millisecond time scale. The problem of determination of intracerebral current sources from a given scalp surface potential is a so called inverse problem with no unique solution. It is necessary to make additional assumptions in order to choose a distinct three-dimensional source distribution among the infinite set of different possible solutions. Regularization methods are:

- *Equivalent dipole/dipole layer localization*: Scanning the head volume with the model source until an error function is minimized.
- *Weighted minimum norm*: Among all possible solutions, choosing the one containing the least energy.
- *Low resolution electromagnetic tomography* (Loreta): Assumes that neighbouring neurons are simultaneously and synchronously activated. Its aim is to find out the smoothest of all possible solutions.

High resistance of the skull is responsible for reduced spatial resolution. It has been shown (Cuffin et al., 2000) that best average localization

that can be achieved is approximately 10 mm using a spherical head model consisting of concentric spheres as brain, skull, and scalp.

FUTURE PERSPECTIVES AND CONCLUSIONS

Due to wavelet transform, there exists a great number of wavelet families. Selecting a certain wavelet depends on previous knowledge of the biophysics of brain processes. It would be desirable to build up a wavelet library for different EEG phenomena.

The reason for the use of a great number of EEG channels is to attain maximum spatial resolution of the scalp voltage distribution to improve topographic mapping considering the inverse estimate problem in neural imaging. A further goal might be to attain realistic head models, and to get individual parameters for the size of brain and skull.

References

- Cuffin, B.N., Schomer, D.L., Ives, J.R. & Blume, H. (2001). Experimental tests of EEG source localization accuracy in spherical head models. *Clinical Neurophysiology*, 112, 46–51.
- Ferree, T.C., Luu, P., Russel, G. S. & Tucker, D.M. (2001). Scalp electrode impedance, infection risk, and EEG data quality. *Clinical Neurophysiology*, 112, 536–544.
- Oostenveld, R. & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology*, 112, 713–719.

Rainer Bösel and Sascha Tamm

Related Entries

PSYCHOPHYSIOLOGICAL EQUIPMENT & MEASUREMENTS, EQUIPMENT FOR ASSESSING BASIC PROCESSES

BURNOUT ASSESSMENT

INTRODUCTION

Job burnout is a prolonged response to chronic interpersonal stressors on the job. It has been recognized as an occupational hazard for various people-oriented professions, such as human services, education, and health care. Recently, as other occupations have become more oriented to customer service, and as global economic realities have changed organizations, the phenomenon of burnout has become relevant in these areas as well. Burnout is defined by the three dimensions of exhaustion, cynicism, and inefficacy. The standard measure that is used to assess these three dimensions is the Maslach Burnout Inventory (MBI).

Burnout is a prolonged response to chronic emotional and interpersonal stressors on the job. It is defined by the three dimensions of exhaustion, cynicism, and inefficacy. As a reliably identifiable job stress syndrome, burnout places

the individual stress experience within a larger organizational context of people's relation to their work. Interventions to alleviate burnout and to promote its opposite, engagement with work, can occur at both organizational and personal levels. The social focus of burnout, the solid research basis concerning the syndrome, and its specific ties to the work domain make a distinct and valuable contribution to people's health and well-being.

CONCEPTUALIZATION

Burnout is a psychological syndrome of exhaustion, cynicism, and inefficacy in the workplace. It is an individual stress experience embedded in a context of complex social relationships, and it involves the person's conception of both self and others on the job. Unlike unidimensional models of stress, this multidimensional model

conceptualizes burnout in terms of its three core components.

Exhaustion refers to feelings of being over-extended and depleted of one's emotional and physical resources. Workers feel drained and used up, without any source of replenishment. They lack enough energy to face another day or another person in need. The exhaustion component represents the basic individual stress dimension of burnout.

Cynicism refers to a negative, hostile, or excessively detached response to the job, which often includes a loss of idealism. It usually develops in response to the overload of emotional exhaustion, and is self-protective at first – an emotional buffer of 'detached concern.' But the risk is that the detachment can turn into dehumanization. The cynicism component represents the interpersonal dimension of burnout.

Inefficacy refers to a decline in feelings of competence and productivity at work. People experience a growing sense of inadequacy about their ability to do the job well, and this may result in a self-imposed verdict of failure. The inefficacy component represents the self-evaluation dimension of burnout.

What has been distinctive about burnout is the interpersonal framework of the phenomenon. The centrality of relationships at work – whether it be relationships with clients, colleagues or supervisors – has always been at the heart of descriptions of burnout. These relationships are the source of both emotional strains and rewards, they can be a resource for coping with job stress, and they often bear the brunt of the negative effects of burnout. Thus, if one were to look at burnout out of context, and simply focus on the individual exhaustion component, one would lose sight of the phenomenon entirely.

In this regard, the multidimensional theory is a distinct improvement over prior unidimensional models of burnout because it both incorporates the single dimension (exhaustion), and extends it by adding two other dimensions: response toward others (cynicism) and response toward self (inefficacy). The inclusion of these two dimensions add something over and above the notion of an individual stress response and make burnout much broader than established ideas of occupational stress.

ASSESSMENT

The only measure that assesses all three of the core dimensions is the *Maslach Burnout Inventory* (MBI), so it is considered the standard tool for research in this field (see Maslach et al., 1996 for the most recent edition). There are now three versions of the MBI, designed for use with different occupations. The original version of the MBI was designed for people working in the human services and health care, given that the early research on burnout was conducted within these occupations and focused on the service relationship between provider and recipient. It is now known as the MBI-Human Services Survey (MBI-HSS). A second version of the MBI was developed for use by people working in educational settings (the MBI-Educators Survey, or MBI-ES). In both the HSS and ES forms, the labels for the three dimensions reflected the focus on occupations where workers interacted extensively with other people (clients, patients, students, etc.): emotional exhaustion, depersonalization, and reduced personal accomplishment.

Given the increasing interest in burnout within occupations that are not so clearly people-oriented, a third, general version of the MBI was developed (the MBI-General Survey, or MBI-GS). Here, the three components of the burnout construct are conceptualized in slightly broader terms, with respect to the general job, and not just to the personal relationships that may be a part of that job. Thus, the labels for the three components are: exhaustion, cynicism (a distant attitude toward the job), and reduced professional efficacy. The MBI-GS assesses the same three dimensions as the original measure, using slightly revised items, and maintains a consistent factor structure across a variety of occupations.

The items in the three MBI subscales are written in the form of statements about personal feelings or attitudes (e.g. 'I feel burned out from my work', 'Working all day is really a strain for me'). The items are answered in terms of the frequency with which the respondent experiences these feelings, on a seven-point, fully anchored scale (ranging from 0 = never to 6 = every day). Because such a response format is least similar to the typical format used in other self-report measures of attitudes and feelings, spurious correlations with other measures (due to

similarities of response formats) should be minimized. Furthermore, the explicit anchoring of all seven points on the frequency dimension creates a more standardized response scale, so that the researcher can be fairly certain about the meanings assumed by respondents for each scale value. The MBI has been found to be reliable, valid, and easy to administer.

As a result of international interest in burnout research, the MBI has been translated into many languages. In most countries, the MBI has simply been translated and its psychometric properties taken for granted. However, some language versions, most notably the French, German, and Dutch versions have been extensively studied psychometrically. Generally speaking, foreign language versions of the MBI have similar internal consistencies and show similar factorial and construct validity as the original American version. Moreover, the three-factor structure of the MBI appears to be invariant across different countries.

Despite these similarities in psychometric properties of the MBI measure, there are national differences in the average levels of burnout. For instance, several studies of various European workers have found lower average levels of exhaustion and cynicism, compared to similar North American samples.

CORRELATES OF BURNOUT

The current body of research evidence yields a fairly consistent picture of the burnout phenomenon (see Schaufeli & Enzmann, 1998). Because burnout is a prolonged response to chronic job stressors, it tends to be fairly stable over time. It is an important mediator of the causal link between various job stressors and individual stress outcomes. The exhaustion component of burnout tends to predict the rise of cynicism, while the inefficacy component tends to develop independently.

The primary antecedents of the exhaustion component are work overload and personal conflict at work. A lack of resources to manage job demands also contributes to burnout. The most critical of these resources has been social support among colleagues. Support underscores shared values and a sense of community within the organization, which enhances employees'

sense of efficacy. Another important resource is the opportunity for employees to participate in decisions that affect their work and to exercise control over their contributions.

Of the three burnout components, exhaustion is the closest to an orthodox stress variable, and therefore is more predictive of stress-related physiological health outcomes than the other two components. In terms of mental, as opposed to physical, health, the link with burnout is more complex. Is burnout itself a form of mental illness, or is it a cause of it? Much of this discussion has focused on depression and burnout, and research has demonstrated that the two constructs are indeed distinct: burnout is job-related and situation-specific, as opposed to depression which is general and context-free.

Burnout has been associated with various forms of job withdrawal – absenteeism, intention to leave the job, and actual turnover. However, for people who stay on the job, burnout leads to lower productivity and effectiveness at work. To the extent that burnout diminishes opportunities for satisfying experiences at work, it is associated with decreased job satisfaction and a reduced commitment to the job or the organization.

People who are experiencing burnout can have a negative impact on their colleagues, both by causing greater personal conflict and by disrupting job tasks. Thus, burnout can be 'contagious' and perpetuate itself through informal interactions on the job. There is also some evidence that burnout has a negative 'spillover' effect on people's home life.

Although the bulk of burnout research has focused on the organizational context in which people work, it has also considered a range of personal qualities. Burnout scores tend to be higher for people who have a less 'hardy' personality or a more external locus of control, or who score as 'neurotic' on the five-factor model of personality. People who exhibit Type-A behaviour tend to be more prone to exhaustion. There are few consistent relationships of burnout with demographic characteristics. Although higher age seems to be associated with lower burnout, it is confounded with both years of experience and with survival bias. The only consistent gender difference is a tendency for men to score slightly higher on cynicism.

FUTURE PERSPECTIVES

The extensive research on burnout has consistently found linear relationships of workplace conditions across the full range of the MBI subscales. Just as high levels of personal conflict are associated with high levels of exhaustion, low levels of conflict are strong predictors of low exhaustion. Conversely, high efficacy is associated with supportive personal relationships, the enhancement of sophisticated skills at work and active participation in shared decision making. These patterns indicate that the opposite of burnout is not a neutral state, but a positive one of job engagement. New research is defining engagement in terms of the positive ends of the three dimensions as burnout. Thus, engagement consists of a state of high energy (rather than exhaustion), strong involvement (rather than cynicism), and a sense of efficacy (rather than inefficacy).

One important implication of the burnout-engagement continuum is that strategies to promote engagement may be just as important for burnout prevention as strategies to reduce the risk of burnout. A workplace that is designed to support the positive development of the three core qualities of energy, involvement, and effectiveness should be successful in promoting the well-being and productivity of its employees, and thus the health of the entire organization.

CONCLUSION AND FUTURE PERSPECTIVES

The personal and organizational costs of burnout provide have led to the development of various intervention strategies. Some try to treat burnout after it has occurred, while others focus on how

to prevent burnout by promoting engagement. Intervention may occur on the level of the individual, workgroup, or an entire organization. At each level, the number of people affected by an intervention and the potential for enduring change increases.

The primary emphasis has been on individual strategies to prevent burnout, rather than social or organizational ones, despite the fact that research has found that situational and organizational factors play a bigger role in burnout than individual ones. Also, individual strategies are relatively ineffective in the workplace, where the person has much less control of stressors than in other domains of his or her life. There are both philosophical and pragmatic reasons underlying the predominant focus on the individual, including notions of individual causality and responsibility, and the assumption that it is easier and cheaper to change people instead of organizations. However, any progress in dealing with burnout will depend on the development of strategies that focus on the job context and its impact on the people who work within it.

References

- Maslach, C., Jackson, S.E. & Leiter, M.P. (1996). *Maslach Burnout Inventory Manual* (3rd ed.). Palo Alto, CA: Consulting Psychologists Press.
- Schaufeli W.B & Enzmann D. (1998). *The Burnout Companion to Study & Practice: A Critical Analysis*. Philadelphia, PA: Taylor & Francis.

Christina Maslach

Related Entries

APPLIED FIELDS: CLINICAL, APPLIED FIELDS: HEALTH, CAREGIVER BURDEN

