MANUAL

BURNOUT ASSESSMENT TOOL[1]

Version 2.0 – July 2020

Wilmar Schaufeli,

Hans De Witte

&

Steffie Desart

[1] Citation: Schaufeli, W.B., De Witte, H. & Desart, S. (2020). Manual Burnout Assessment Tool (BAT) – Version 2.0. KU Leuven, Belgium: Unpublished internal report.

Contents

Acknowledgements

Many people and organizations contributed to the development of the Burnout Assessment Tool (BAT). Without their active input this manual could not have been written. We would like to thank our partner organizations for the fruitful cooperation in developing the BAT. First and foremost, we would like to thank the Idewe Group, Pulso Europe, ArboNed, The Human Link, Compass, and Work Life Balance for their input in the discussions about the definition of burnout. Moreover, we would like to thank Securex and Better Minds at Work for data collecting among employees in Flanders, and Burn-out.Flanders, Bye Bye Burnout, U-count, Compass, ArboNed and ArboUnie for data collection among those who suffer from burnout.

We also thank Ellen Caers, Kaat Vanbrabant, Dries de Bie, Jana Wouters and Chiel Buseyne who have made an important contribution to this manual with their master's thesis, and Ewa Gruszczyńka for het assistance with the Latent Class Analysis.

We would like particularly to thank the many thousands of anonymous respondents who have taken the time to complete an extensive questionnaire. This enabled us to write this manual, which introduces a new burn-out instrument for Flanders and the Netherlands.

The development of the BAT was sponsored by a grant from KU Leuven[2].

[2] Project code: C32/15/003 – Development and validation of a questionnaire to assess burnout

Abstract

The Burnout Assessment Tool (BAT) is a new self-report questionnaire to measure burnout. Its development is deemed necessary because the most frequently used questionnaire – the Maslach Burnout Inventory (MBI) – shows a number of conceptual, technical and practical imperfections. As such this is not surprising because the MBI was developed almost forty years ago for research and not as an assessment tool.

The BAT was developed using a combination of a deductive (theoretical) and an inductive (empirical) approach. Theoretically, burnout is seen as a state of mental exhaustion that manifests itself as both the inability and unwillingness to spend effort at work. Both are two sides of the same coin; on the one hand the inability to carry out one's work properly because of chronic fatigue ("I can no longer do my job"), and on the other hand the unwillingness to perform because of mental distance towards one's job ("I do not want to do my job anymore").

On the basis of a number of in-depth interviews with professionals who deal with burned-out employees on a daily basis, two more aspects were added; emotional and cognitive impairment. That means that the functional capacity to regulate one's emotional and cognitive processes is impeded, which manifests itself, for example, in sudden bursts of anger and in poor concentration. In addition, psychological stress symptoms, psychosomatic complaints and depressed mood were mentioned by the professionals, which can be considered as secondary burnout symptoms. These complaints are "atypical" because they do not only occur in people who have burnout. However, because people who suffer from burnout often exhibit such symptoms, it is important to take these into account as well.

In addition to in-depth interviews, an analysis was carried out of over a dozen burnout questionnaires that included more than 300 items. It appeared that exhaustion plays a pivotal role in *all* questionnaires and that mental distance is included in all questionnaires that do not limit burnout to mere exhaustion.

Based on the interviews and the analysis of the questionnaires, four scales are proposed: (1) exhaustion (8 items), mental distance (5 items,) emotional impairment (5 items), and cognitive impairment (5 items). In addition, a general version of the BAT has also been constructed that can be filled out by those who have not been working for a longer time, for instance, because they are sick listed.

Psychometric research in two representative samples of the Flemish and Dutch workforce shows that these four aspects of burnout can indeed be distinguished. However, they are strongly interrelated and refer to one underlying condition. It also appears that 6 groups can be distinguished that differ in levels of burnout on each of the four subscales. These findings are consistent with the notion that burnout is a *syndrome* of four interdependent dimensions. This implicates that the total score of the BAT can be used as an indicator for burnout.

The atypical psychological and psychosomatic symptoms appear to be so strongly interrelated that they constitute one scale, which can be distinguished from depressed mood. A single weak item was removed, leaving 10 secondary symptoms for the entire scale. No separate scale for depressed mood was developed, because many valid and reliable depression scales already exist.

The internal consistency and test-retest reliability of the BAT is good to excellent, but inter-rater reliability is weak to moderate, especially of the subscales. The BAT shows the expected similarity with other burnout questionnaires, including the MBI (convergent validity), but at the same time it also contributes independently to these burnout questionnaires (divergent validity). Furthermore, burnout, as measured by the BAT, can be distinguished from work engagement, boredom, and workaholism. The core symptoms of the BAT cannot be distinguished from the secondary symptoms, which suggests that both refer to the same underlying psychological state. In short, the BAT measures burnout and no other related concepts (discriminant validity).

The content validity of the BAT is confirmed by the fact that its subscales, as well as its total-score are, as expected, positively associated with job demands, such as role conflict and bureaucracy and negatively associated with job resources, such as person-job fit and team atmosphere. BAT-scores are also related to organizational outcomes, such as turnover intention and absenteeism, and with personality traits such

as neuroticism. The content validity of the general version of the BAT is supported by the expected associations with a number of well-being indicators, including subjective health, happiness and depression.

Statistical norms have been calculated on the basis of percentile scores, so that BAT scores can be classified as "low", "average", "high" and "very high" relative to the Flemish and Dutch workforce. Furthermore, based on a group that suffers from serious burnout cut-off scores have been calculated. These can used to identify employees who are at risk of burnout or are most likely suffering from serious burnout. To date, such cut-off values are only available for Flanders.

The BAT can be used individually for burnout assessment as well collectively in organizations for screening and identifying those with elevated burnout levels. Also, the BAT can also be used for benchmarking; that is, for determining the level of burnout relative to a representative sample of the workforce. Finally, the general version of the BAT can be used for assessing and monitoring those who are currently not working, for example, in the context of a treatment, counseling, or return-to-work program.

In addition to the full version of the BAT with 23 items, a shortened version with 12 items is also available. The items of the shortened version were selected using Rasch analysis. The psychometric properties of the BAT-12 (i.e., factorial validity, internal consistency, test-retest reliability, and convergent and discriminant validity) are similar to that of the full version. This also applies to the relationship with job demands, job resources, organizational outcomes and personality traits (content validity). Furthermore, the corresponding subscales and the total scores of the two versions of the BAT correlate almost perfectly with each other. The BAT-23 is recommended for individual diagnostic assessment, whereas the BAT-12 can be used just as well for other purposes, such as screening and monitoring.

PART I: INTRODUCTION

1. Concept and measurement of burnout

Burnout is a metaphor that refers to a state of mental exhaustion. It was first used in the United States at the end of the 1970s (Maslach & Schaufeli, 1993). However, various examples of burnout *avant-la-lettre* suggest that mental exhaustion is of all times (Schaffner, 2016). Particularly interesting is the parallel with neurasthenia – literally nervous weakness (Schaufeli, 2017a), which emerged at the end of the 19th century and was seen as the consequence of modern, hectic life. It was argued that sensory overstimulation by talking on the telephone, reading newspapers and traveling by steam train would weaken the nerves, resulting in extreme fatigue and work incapacity. The similarity with burnout is striking, this too is seen as a modernity ailment that is the result of overstimulation, but currently due to social media, and psychosocial demanding (work) life. The result is the same: mental exhaustion.

Burnout has become a social and scientific issue that has been the focus of much attention for almost four decades. It seems that our work has become more and more stressful, as a result of which a growing number of individuals feel exhausted and drop out of work due to illness or disability. So far we neither know the exact prevalence of burnout nor how much it increased over time (Schaufeli, 2018).

Yet, there are national and European laws and regulations that oblige employers to periodically assess psychosocial risks among their employees and to implement policies to prevent burnout and jobstress[3]. In a number of European countries, including Belgium and the Netherlands, burnout is recognized as an occupational disease or work-related disorder (Lastovkova, Carder, Rasmussen, Sjoberg et al., 2017).

[3] See: https://www.eurofound.europa.eu/publications/report/2018/burnout-in-the-workplace-a-review-of-data-and-policy-responses-in-the-eu

1.1. What is burnout?

What exactly are we talking about when we speak of burnout? In everyday language and in the media, the term is usually used loosely and refers to "not feeling well at work". In the context of occupational, mental, and psychosocial care, however, burnout refers to a mental disorder that is characterized by severe mental exhaustion among those who cannot cope without the help or assistance of others. Often there are also other complaints and besides work, the private situation also plays a role. In science, burnout is generally regarded as a multidimensional concept with exhaustion as its central feature. Although different definitions of burnout exist, scholars consider exhaustion as its core characteristic.

Initially, burnout was seen as something that exclusively occurs in the human services, such as in social work, health care and education, but also in law enforcement and among poverty lawyers. Working with people requires quite some energy, effort, attention and empathy and is therefore emotionally demanding. It was reasoned that inter-personal stress accumulates and when recovery is insufficient, this may ultimately lead to mental exhaustion (Schaufeli, Leiter & Maslach, 2009).

The most commonly used definition that stems from the early period of burnout research is in line with this reasoning and comes from Maslach and Jackson (1981a; p. 99): "Burnout is a syndrome of emotional exhaustion and cynicism that occurs frequently among individuals who do 'people work' or some kind". In addition to emotional exhaustion and a negative, cynical attitude towards recipients, a third factor was distinguished: reduced personal accomplishment, or the tendency to judge oneself negatively in contact with recipients with whom one works professionally. It is important to realize that initially – by definition – burnout could only occur among those who work with other people. This implied that those who didn't could not suffer from burnout!

1.2. The Maslach Burnout Inventory

Based on the above definition, the Maslach Burnout Inventory (Maslach & Jackson, 1981b) was developed, a self-assessment questionnaire that includes emotional exhaustion ("feelings of being emotionally overextended and exhausted by one's

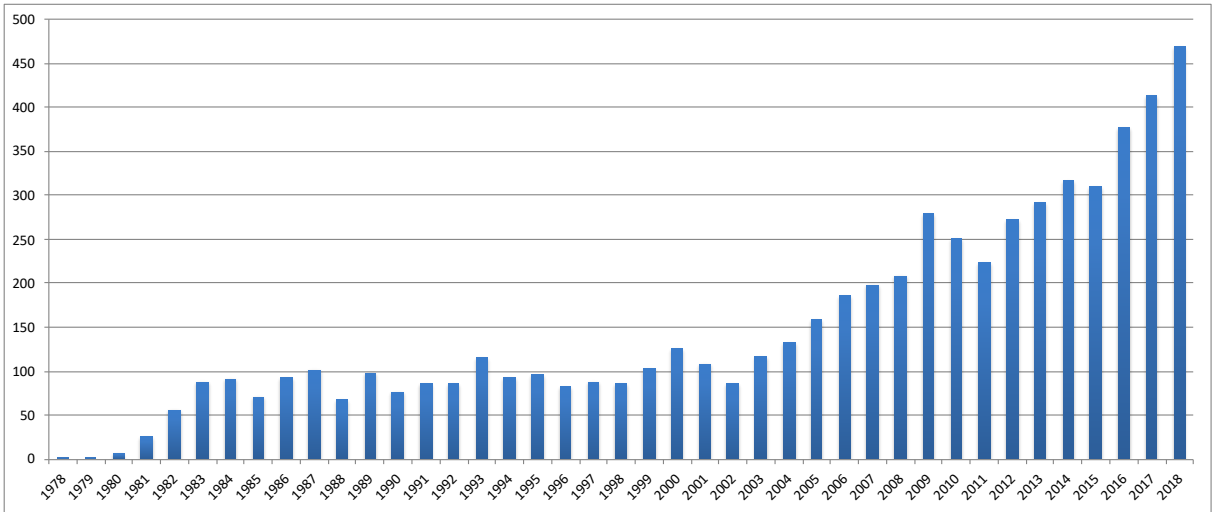
work"), depersonalization ("an unfeeling and impersonal response towards recipients of one's service, care, treatment, or instruction") and personal accomplishment ("feelings of competence and successful achievement in one's work with people"). A high score on emotional exhaustion and depersonalization and a low score on personal accomplishment was supposed to be indicative for burnout. In addition to a version for care providers in the broad sense (MBI-HSS - Human Services Survey), an almost identical version for teachers (MBI-ES - Educators Survey) was introduced in the 2nd edition of the MBI test manual (Maslach & Jackson, 1986), in which essentially "recipients" was replaced by "students".

The 3rd edition of the test manual (Maslach, Jackson & Leiter, 1996) is a milestone because the MBI - General Survey (GS) was introduced. This questionnaire makes it possible to assess burnout among *all* employees, including those who do not do "people work of some kind". The MBI-GS defines burnout in a more general sense as "a state of exhaustion in which one is cynical about the value of one's occupation and doubtful or one's capacity to perform" (p. 20). In contrast to the MBI-HSS and MBI-ES, the questions about exhaustion of the MBI-GS do not refer to other people as the cause of fatigue. Similarly, the questions about cynicism refer to the work itself and not to the relationships with others at work. And finally, the questions about professional efficacy are formulated more broadly and also refer to non-social aspects of one's work. Recently, the 4th edition of the manual was published (Maslach, Jackson, Leiter, Schaufeli & Schwab, 2016) that incluses versions of the MBI for medical staff and students.

Although different self-report questionnaires exist to measure burnout, the MBI has been the gold standard from the start. In 1998, Schaufeli and Enzmann found that the MBI was used in 91% of all scientific publications on burnout. It appears that after almost twenty years not much has changed because Boudreau, Boudreau and Mauthe-Kaddoura (2015) estimated that the MBI was still used in 88% of tall publications on burnout. In fact, that means that burnout is what the MBI measures, and vice versa. This mutual dependence of object and method – in combination with the dominance of the MBI – is undesirable because this impedes new and innovative research that can lead to a better understanding of burnout.

Meanwhile, scientific interest in burnout has increased rapidly, especially after the turn of the century, as shown in figure 1. This increase seems to correspond with the availability of the MBI-GS, which enabled the study of burnout among all kinds of workers.

Figure 1: Increase in the number of scientific publications on burnout 1978-2018[4]



Source: PsycINFO (Juy, 2020)

This large body of scientific publications has increased our knowledge about burnout, but there are still important gaps, especially when it comes to applied knowledge on epidemiology, assessment, prevention, treatment, and return to work (Schaufeli, 2018). Tellingly, it is precisely in these areas where the MBI – known in the Netherlands as the Utrecht Burnout Scale (UBOS: Schaufeli & Van Dierendonck, 2000) – fails to deliver. This is not surprising because at the time the MBI was developed for scientific research purposes and not for practical use in screening, assessment, and monitoring prevention and treatment programs. Below, the problems with the MBI are further elaborated, which illustrates the need to develop a new burnout questionnaire.

[4] Figure 1 only includes publications from English-language, peer-reviewed psychological journals with 'burnout' in the title. In total this amounts to 6.149 articles. In an approximately 12.700 articles burnout, is mentioned as a keyword. Google Scholar includes more than 109,000 publications on burnout, this includes books, book chapters, research reports and non-English-language publications as well.

2. Problems with the measurement of burnout

Because the MBI is considered the gold standard, to measure burnout we limit ourselves to discussing the problems with the MBI and do not discuss the limitations of other burnout questionnaires. The problems with the MBI are conceptual, technical and practical in nature.

2.1. Conceptual problems

After interviewing burnout victims Maslach and Jackson (1981a, b) formulated 47 items which, in their view, would tap burnout. This item pool was presented to more than 600 human services professionals. After factor analysis, ten factors emerged, four of which explained the lion's share of the variance. In addition to the three dimensions mentioned above, a fourth dimension emerged: lack of involvement. Initially, this was considered an optional dimension that was gradually neglected though. The number of items was then reduced to 25 and presented to a second, similar sample of more than 400 professionals. Because the results did not deviate in this second group, both samples were pooled into a total group of more than 1,000 professionals that was used as the basis for the first version of the MBI manual. It is important to mention that this is *not* a representative sample of American human service professionals. That is the very reason why there are no useful test norms available for the MBI (see below).

Hence, an inductive, bottom-up approach was used to develop the MBI. The test-authors formulated questionnaire items based on oral and written information that they had received from members of a particular occupational group. Please note that no people with serious burnout symptoms were sampled, instead those were included who worked in professions with a supposedly high burnout risk.

Because an inductive approach has been used, there is some arbitrariness about the number and nature of the dimensions of burnout in the MBI. For example, it has now been established that burnout is also characterized by reduced cognitive performance, meaning that cognitive functions such as attention, concentration and working memory are impaired (Deligkaris, Panagopoulos, Montgomery & Masoura, 2014). However, the MBI does not include this cognitive dimension, so that it is no longer up-to-date.

Furthermore, serious doubts has been risen as to whether reduced personal accomplishment or professional efficacy is a constituting element of burnout (Schaufeli & Taris, 2005). This may be interpreted alternatively as a cause or consequence of burnout. In the former case exhaustion would occur when a personal resource such as personal accomplishment or professional efficacy is lacking, whereas in the in the latter case feelings of reduced accomplishment or efficacy may result from poor performance that is caused by mental exhaustion (see further under 3.2).

2.2. Technical problems

Because of the extreme wording the answers of some items are very skewed, which, in turn, may affect the reliability of the MBI (Wheeler, Vassar, Worley, & Barnes, 2011). This is especially true for depersonalization items such as "I don't really care what happens to some recipients" and "I feel that I treat some recipients as if there were impersonal objects"[5]. The overwhelming majority of people in Flanders and the Netherlands indicate that this "never" or only "sporadically" applies, while virtually no one ticks "very often" or "always".

In the original American version of the MBI, fixed response categories (anchors) are used, such as "a few times a year or less" (score 1) and "once a month or less" (score 2). However, the problem is that the distance between these response categories is not equivalent. For instance, to the distance between scores 1 and 2, is not the same as between scores 2 and 3 ("a few times a month"). In the Dutch version an attempt was made to solve this problem by using variable anchors in addition to fixed anchors, such as "sporadic" (score 1), "occasional" (score 2), and "regular" (score 3). This is not ideal because it may lead to confusion because both scoring methods may be interpreted differently by the respondents. In addition to the nature of the response categories, the 7-point answer scale (score 0-6) of the MBI is highly differentiated and one may wonder whether people are capable of making a clear distinction between, for example, "sporadic" and "occasionally".

[5] In the past when only paper-and-pencil versions of the MBI were used, it often happened that respondents spontaneously added comments on the answering sheet such as, "Of course not !!" and "What nonsense !!"

A third and final psychometric problem with the MBI is that the questions about personal accomplishment or professional efficacy are positively worded; that is, a *low* rather than a high score is indicative for burnout. As a consequence, this dimension usually behaves differently compared to the other two burnout dimensions. For example, the correlation of this positively worded dimension with the other two dimensions is usually much lower than the interrelation between these two dimensions. This is illustrated by the norming sample of the UBOS (N = 11,429), where the correlation between personal accomplishment on the one hand and exhaustion and depersonalization on the other is -20 and -.30, respectively, compared to .49 between exhaustion and depersonalization (Schaufeli & Van Dierendonck, 2000; p. 21). The same pattern is found in a meta-analysis of 45 studies, which shows that exhaustion and depersonalization correlate on average .56, while their association with personal accomplishment is much less; r = -.30 with exhaustion and r = -.35 with depersonalization (Worley, Wheeler & Barnes, 2008).

The special role of personal accomplishment is also reflected in a study by De Beer and Bianchi (2019) that demonstrates that a two-factor model consisting of a *combined* exhaustion and depersonalization factor and a separate personal accomplishment factor fits best with the data. This corresponds with an earlier finding by Schaufeli and Bakker (2004a) that the efficacy items of the UBOS constitute a positive factor together with the three subscales of work engagement (vigor, dedication and absorption), while exhaustion and cynicism items constitute a second, separate, negative burnout factor.

Moreover, it is striking that when the questions are reworded so that they refer to *in*competence instead of competence, correlations with the other two burnout dimensions are much higher – and positive (Bresó, Salanova & Schaufeli, 2007). So it seems that we are dealing with a statistical artifact that is caused by using positive instead of negative items to measure burnout; as a consequence, the negatively formulated accomplishment or efficacy items are more strongly related to both other MBI dimensions than when reversed, positively worded items are used. Below we will see that there are also other doubts as to whether reduced personal accomplishment or professional efficacy is a constituting element of burnout (see 3.2). All in all, the third burnout dimension is a stranger in our midst.

2.3. Practical problems

The practical use of the MBI is hampered by the fact that it does not produce a single burnout score but three separate scores for each subscale. The MBI test manual explicitly states: "In general, each respondent's scale scores should be calculated and interpreted separately. Note that responses to MBI items should not be combined to form a single 'burnout' score" (Maslach et al., 2017; p. 44). This once again illustrates that the MBI was never intended as a diagnostic tool to identify people with severe burnout complaints, but rather to investigate – on different dimensions – the relationships of people with their jobs. The fact that the MBI yields three scores instead of one creates a certain ambiguity because the MBI manual also refers to burnout as a "syndrome". This implies that the three dimensions are closely interrelated and refer to one underlying psychological state or disorder (i.e., burnout). However, a Flemish study by Van Heule, Rosseel, Vlerick, Van de Ven and Declercq (2012) shows that a second-order factor model, in which all three components of the UBOS load on one common, underlying factor, does not fit with the data. This second-order model represents burnout as a syndrome with three dimensions that each constitute an integral part of the underlying burnout condition.

In order to meet the need of practitioners for a single burnout score, the UBOS test manual recommends using the so-called one-plus-two rule, following Brenninkmeijer and Van Yperen (2003). This rule states that a person is likely to suffer from burnout when he or she exhibits a (very) high score on exhaustion in combination with *either* a (very) high score on depersonalization (or cynicism), *or* a (very) low score on personal accomplishment (or professional efficacy).

Clinically validated cut-off scores for MBI do not exist in most countries. The Netherlands is one of the very few exceptions – if not the only one. Such cut-off values can be used to identify people with severe burnout and are based on a group of individuals who, according to experts such as psychologists, general practitioners or occupational physicians, actually suffer from burnout (Schaufeli & Van Dierendonck, 2000; pp. 43-51; Schaufeli, Bakker, Schaap, Kladler & Hoogduin, 2001). However, a replication study by Kleijweg, Verbraak and Van Dijk (2013) on the clinical validity of the UBOS among 419 clients of a treatment center for work-related psychological

problems failed to confirm its discriminatory power. The UBOS was unable to distinguish between those who did and did *not* receive the DSM diagnosis "undifferentiated somatoform disorder" that was used as a proxy for burnout. The authors concluded: "The practical implication is that the MBI should not be used by itself as a diagnostic tool in a patient population, because of a resultant high probability or overdiagnosing burnout" (p. 435). In addition, the norms of the Dutch version of the MBI are now outdated and should no longer be used according to the guidelines of the Netherlands Institute of Psychologists (NIP).

Previous manuals of the American MBI included statistical norms for each of the three dimensions. These were based on the upper and lower terciles of the score distribution of the overall reference group, which, by the way, was not representative, as noted above. That means that someone who is classified as "high" or "low" on a particular dimension scores among the 33.3% highest, or lowest scorers, respectively. Accordingly, it is assumed that one third of the reference group – namely those scoring "high" – suffer from burnout. Whether this is indeed the case is shrouded in mystery. This is the very reason why these statistical norms are no longer included in the most recent version of the MBI manual (Maslach et al., 2017). Instead, the interpretation of the scores is now based on either the absolute score or on a comparison with the average of a particular reference group from the manual. In the former case this means that an exhaustion score of, for example, 3.5 indicates that someone feels on average "a few times a month" exhausted. In the latter case, a score of 3.5 means that somebody scores above the average of a non-representative sample. In short, both methods from the American manual are inappropriate for an accurate individual assessment of burnout because a comparison with a group that actually suffers from burnout is not possible.

3. Basic principles of the BAT

Our starting point was the idea that the BAT should avoid the aforementioned conceptual, technical and practical problems that are associated with the MBI. In addition, a version of the BAT should be available that can be used among those who are not working, for instance because they are on sick leave because of burnout (see 5 below). The following five principles have been used for developing the BAT.

Principle 1. *Combining deductive and an inductive approaches*

Our deductive approach is based on the principle that the BAT should be in line with the theoretical description of burnout, as provided by Schaufeli and Taris (2005). They followed Edward Thorndike (1874-1949), the grand old man of psychological fatigue research, who maintained that the basic tenet of fatigue is "the intolerance of any effort". In his view, fatigue is both the inability *and* the unwillingness to spend effort, which is reflected by its energetic and motivational component, respectively. The unwillingness to perform manifests itself by increased resistance, reduced commitment, lack of interest, disengagement, and so on – in short, mental distancing. Thus, according to Schaufeli and Taris (2005), inability (exhaustion) and unwillingness (distancing) constitute two sides of the same burnout coin.

As pointed out above, burnout has an energetic and a motivational component, reflected by no longer being *able* to work and no longer *wanting* to spend effort at work. Meijman and Schaufeli (1996) argued that the unwillingness to spend any more effort is functional in acute psychological fatigue. For example, the resistance to spend effort at work grows in the course of the working day as the employee becomes increasingly tired. This declining motivation indicates that it is time to stop working, so that one can recover. So in fact, this decreasing motivation in the form of mental distancing acts as a protective mechanism to prevent the individual from spending additional energy and thus entirely depleting his or her resources. In "normal" occupational fatigue, distancing and withdrawal is functional because it fosters recuperation or switching to another task, and will therefore reduce fatigue. In chronic fatigue, however, the protective distancing mechanism is dysfunctional because it has habituated into relatively permanent impaired motivation. Mental distancing has become a part of occupational life: instead of a solution it is now part of the problem that we call burnout. Mental distancing can therefore be seen as an inadequate coping strategy that promotes rather than reduces exhaustion, as is illustrated by a longitudinal study among teachers and nurses (Taris, Le Blanc, Schaufeli & Schreurs, 2005). So much for the deductive approach.

In addition we used an inductive approach is in the first phase of the development of the BAT. On the one hand we conducted a number of in-depth interviews with practitioners from Flanders and the Netherlands, and on the other hand we performed

a literature study and item analysis of questionnaires that claim to assess burnout. The end result of the first endeavor was a list of burnout symptoms and of the second a list of questionnaire items. The clustering of symptoms and items was guided by the description of burnout as combination of the inability and unwillingness to spend effort at work.

In sum, that the BAT was developed on the basis of the dialectic between an a priori, theoretical description of burnout (deduction) as well as an empirical inventory of burnout characteristics (induction). This approach has resulted in a new definition of burnout (see 4.3).

Principle 2. *Up-to-date content*

Burnout is not only characterized by exhaustion and mental distance, but also by impaired cognitive functioning, such as poor attention and concentration, and a poor working memory (Deligkaris, et al., 2014). Moreover, such cognitive impairments are particularly persistent. Although two years after treatment burnout patients performed better on a number of cognitive tasks, their performance level was still less than that of normal, healthy subjects (Van Dam, Keijsers, Eling & Becker, 2012). A comparable Swedish study draws a similar conclusion; even three years after patients with burnout sought help, their cognitive functions that are related to speed, attention and memory are still not at the normal, healthy level (Johnsdottir, Noldlund, et al., 2017). So it seems that cognitive impairment is not only part of the burnout syndrome, but also lasts for a very long time, even after other burnout symptoms have diminished. Taken together, this is reason enough to specifically scan the list of symptoms and items, that results from the first phase of the development of the BAT, for indications of impaired cognitive functioning.

Furthermore, reduced personal accomplishment or professional efficacy is not included in the BAT for conceptual and empirical reasons. As argued above, inability and unwillingness to spend effort constitute the core of burnout. Reduced personal accomplishment or professional efficacy refers to the tendency to evaluate oneself negatively with regards to the work that one performs. Conceptually, this comes very close to self-efficacy; the individual's belief in their innate ability to achieve goals (Bandura, 1997). This is illustrated by the relatively high correlation of .49 between the

two concepts found in a meta-analysis of 57 studies (Shoji, Cieslak, et al., 2016). Contrarily, the correlation of self-efficacy with exhaustion and depersonalization is significantly lower – ρ = -.31 and ρ = -.33, respectively. As such, personal accomplishment or professional efficacy can be seen as a consequence of being unable and unwilling to no longer spend effort at work. After all, because people cannot and do not ant longer want to perform at work, they feel inadequate and incompetent. This view is supported by two longitudinal studies conducted by Taris, et al. (2005) among Dutch teachers and nurses. These studies showed that exhaustion leads to depersonalization and depersonalization leads, in its turn, to reduced personal accomplishment. It is also important to note that reduced accomplishment or efficacy was virtually never mentioned in the in-depth interviews with practitioners (see 4.1 below). This again suggests that this component of the MBI plays a minor or no role at all in burnout.

Principle 3. *Carefully chosen item formulations and response categories*

When formulating the BAT items, care has been taken to ensure that they are easy to understand and as short and concise as possible. Furthermore, strongly formulated questions are avoided because these might result in skewed answers.

Only items are included in the BAT that directly refer to burnout symptoms. In order to avoid a possible artifact such as with the MBI (see 2 above), no items are used that need to be "reversed". This agrees with Hartley (2013), who concluded from an overview of response scales in clinical and health psychology; "It is best either (a) to remove negatively worded items from a scale, or (b) to present the results for such items separately'"(p. 84).

A 5-point Likert frequency scale with the following values is used as answering format for the BAT: 1 (never), 2 (rare), 3 (sometimes), 4 (often) and 5 (always). Experience from practice shows that respondents feel more confident to make a distinction between 5 rather than between 7 or more response categories.

Principle 4. *Diagnostic use*

Because of its use as an assessment tool, the BAT should provide a single burnout score. This means that the BAT assumes that burnout is a syndrome that consists of interrelated

symptoms that all refer to one underlying psychological condition. This does not alter the fact that burnout symptoms or dimensions may also be used separately, for example, to further specify the general picture, which may be particularly important for individual burnout assessment.

Diagnostic use also implies that clinically validated cutoff values should be available for the BAT. These must be based on a well-defined group of people who meet specific inclusion criteria for burnout. In other words, cut-off values of the BAT are established on the basis of the scores of people who unambiguously suffer from burnout. By using these as a point of reference, the individual's risk on burnout can be assessed.

Principle 5. *General version*

Although burnout is a work-related phenomenon, a need for a context-free instrument exists for those who currently do not work. After all, someone who does not work cannot answer an item such as: "I feel empty at the end of a working day." Nevertheless, the MBI that includes this particular item, is often wrongly administered among employees who are on long-term sick leave (e.g., De Vente, Olff, Van Amsterdam, Kamphuis & Emmelkamp, 2003). Instead, of course, the item should read: "I feel empty at the end of the day." The assumption is that exhaustion, although it originates from work, may still be present, even after one has stopped working. In a similar vein, mental distance to one's work may exist after stopping with work.

4. The development of the BAT

The development of the BAT took place in four consecutive steps. In the *first step*, in-depth interviews were conducted with professionals from Flanders and the Netherlands and an inventory was made of burnout questionnaires. Based on this information, a provisional version of the BAT was composed. In the *second step*, a number of pilot studies were conducted that focused on the scoring categories and different versions of the BAT (work-related vs. general). In the *third step*, the psychometric qualities of the final version of the BAT were extensively investigated in two large, representative Dutch and Flemish samples. And finally, in the *fourth step*, norms and cut-off values were calculated, using a sample of employees who actually

suffer from serious burnout. The first two steps are described below and the third and fourth steps are discussed in detail in the remaining parts of the manual.

4.1. In-depth interviews

In-depth, structured interviews of 1-2 hours were conducted with Flemish and Dutch professionals who frequently deal with persons who suffer from burnout. The aim was to find out which symptoms practitioners identify as typical for burnout. A total of 49 interviews were held with 19 general practitioners, 17 occupational physicians and 13 psychologists[6]. Of the psychologists, about one half focused on burnout prevention in organizations, whereas the other half counseled or treated burned-out employees in a clinical setting as a coach or psychotherapist. Hence, a heterogeneous group of professionals was interviewed in order to obtain a broad picture on burnout.

Interviewing these professionals also provided us with a deeper understanding of the nature and background of burnout. First, the interviewees were invited to describe a typical burnout case. On the basis of that case, the specific symptoms and causes were discussed in and the professionals were invited to offer their own definition of burnout. Subsequently, the symptoms that were mentioned were ranked according to their importance for assessing burnout. Attention was also given to a-typical symptoms that are observed in other psychological disorders as well and to symptoms that may discriminate burnout from depression.

The qualitative data from the interviews were analyzed with the Nvivo program. This *Computer Assisted Qualitative Data Analysis* program allows to organize the data and prepare it for further analyzes. By performing an inductive content analysis particular aspects can be clustered, such as symptoms that emerged from interviews (Elo & Kyngäs, 2008).

---

[6] Thanks to Ellen Caers and Kaat Vanbrabant, who both wrote their master's thesis in the context of this research at KU Leuven.

- Caers, El. (2017). *Are the occupational physicians and science aligned? A comparative study between burnout in science and in practice*.
- Vanbrabant, K. (2017). *Burnout: Does the GP know what he's talking about? An investigation into the correspondence between literature and practice*.

A total of 260 different symptoms were mentioned by the practitioners. These symptoms could be summarized in 19 different clusters in the first round. In second the round this could be further reduced to seven dimensions which, on the basis of theoretical considerations (see 3.1), were subdivided into four core aspects and three secondary aspects of burnout. Not surprisingly, fatigue was mentioned unanimously as the most important burnout symptom (e.g., "exhaustion", "feeling empty", "completely exhausted", "lacking any energy", and "looking tired"). In addition, symptoms emerged that refer to cognitive and emotional impairment. Examples of the former are: "concentration problems", "making mistakes", "disturbed imprinting", "being less efficient", and "forgetfulness"; and of the latter: "weeping," "irritability," "anger", "hot temper," and "being emotional." Finally, symptoms of mental distance were mentioned, such as "no motivation," "withdrawal," "finding the work meaningless", "indifference," and "cynicism." So far the four core or primary symptoms of burnout.

In addition, depressive complaints were mentioned, such as "gloominess", "feeling of being worthless", "loss of initiative", and "apathetic", as well as psychological distress such as "anxiety", "sleep problems", "worrying", and "nervousness", and psychosomatic complaints, such as "headache", "stomach and bowel problems", "muscle weakness", and "palpitations". These three types of complaints are considered secondary and atypical, because they may also occur with other physical and psychological disorders, such as cancer, hypo- or hyperthyroidism, mood disorder, or anxiety disorder.

4.2. Inventory and analysis of burnout questionnaires

To obtain an overview of the items that are used to measure burnout, an inventory was made of self-report questionnaires. As indicated earlier, the MBI is by far the most frequently used measure (see 1.1). Nevertheless other questionnaires are used as well. Our literature search yielded 11 other self-reporting questionnaires that were divided into four groups based on their psychometric quality (reliability and validity) and frequency and nature of their use.

The first group of nine questionnaires has sufficient psychometric quality:

1. Maslach Burnout Inventory (MBI; Maslach, & Jackson 1981a, b; see also 1.1);

2. Oldenburg Burnout Inventory (OLBI; Demerouti, Bakker, Vardakou & Kantas, 2003);
3. Bergen Burnout Inventory (BBI; Salmela-Aro, Rantanen, Hyvönen, Tilleman, & Feldt, 2011);
4. Copenhagen Psychosocial Questionnaire (COPSOQ; Kristensen, Borritz, Villadsen, & Christensen, 2005);
5. Spanish Burnout Inventory (Gil-Monte & Faúndez, 2011);
6. Granada Burnout Questionnaire (GBQ; De la Fuente, et al., 2013);
7. BurnOut-Neuratshenia Complaints Scale (BONKS; Verbraak, Van de Griendt & Hoogduin, 2006);
8. Shirom Melamed Burnout Measure (SMBM: Shirom & Melamed, 2006);
9. Burnout Measure (BM; Pines & Aronson, 1981).

The psychometric qualities of the second group of three questionnaires are unknown:

1. Boudreau Burnout Questionnaire (BBQ; Boudreau et al., 2006);
2. Hamburg Burnout Inventory (HBI, Burisch, 2017);
3. Instrument for the early detection of burnout (IVOB; Federal Public Service, 2017).

The questionnaires from these two groups were used for further analysis. Other burnout questionnaires were not included because they related to specific professions, such as doctors (Physician Burnout Questionnaire; Morenzo-Jimenez, Barbaranelli, Galvez-Herre, & Garrossa, 2012) or teachers (Teacher Burnout Scale; Seidman & Zager, 1987). Finally, the Four-Dimensional Complaints List (4DKL; Terluin, 1996) was selected for further analysis because it is frequently used by occupational health professionals in the Netherlands (Terluin, Van Rhenen, Schaufeli & De Haan, 2006). In fact, this questionnaire does not measure burnout per se, but four types of mental complaints that are most common in primary health care: somatization, distress, anxiety, and depression. In the Netherlands, the 4DKL is also used to assess over-strain ("surmenage"), a mental state that largely overlaps with burnout, albeit that feelings of exhaustion are less prominent and the symptoms are less serious and persistent (Verschuren, Nauta et al., 2011). According to the Dutch multidisciplinary guideline on

job stress and burnout for primary care professionals[7], over-strain lasts less than 6 months and burnout more than 6 months.

A total of 13 questionnaires were analyzed, which included 327 items that were grouped into 50 different subscales. On average, each questionnaire contains three to four subscales with the "exhaustion" scale being included without exception in every questionnaire, usually as mental exhaustion but sometimes also in combination with physical exhaustion. Three questionnaires restrict burnout exclusively to exhaustion (COPSOQ, BM and SMBM), while two questionnaires add secondary symptoms (BONKS and IVOP), and one questionnaire only consists of secondary symptoms (4DKL). All remaining seven questionnaires that measure burnout as a multi-dimensional concept include a scale that taps mental distance.

Five questionnaires use a 5-point Likert answering scale, four use a 7-point scale, and another four use a 4-point or a 3-point scale. The vast majority of the items have been formulated negatively; only 4% of all subscales consist of exclusively positively worded items, and 20% consist of mixed scales with both positively and negatively worded items.

To conclude; agreement exists about exhaustion as the core dimension of burnout. Furthermore, all multi-dimensional burnout questionnaires include – in addition to exhaustion – a subscale that measures mental distance. In other words; the content of all multi-dimensional burnout questionnaires reflects the conceptualization of burnout presented earlier as a combination of inability (exhaustion) and unwillingness (mental distance) to perform.

4.3. The conceptualization of burnout of the BAT[8]

Based on the in-depth interviews and analysis of burnout questionnaires, and against the background of the theoretical considerations about the nature of burnout (see

[7] https://www.nvab-online.nl/sites/default/files/bestanden-webpaginas/MDRL_Overspanning-Burnout.pdf

[8] This conceptualization is also included in the report on burnout and work of the Belgian Superior Health Council (September 2017, HGR no. 9339); see https://www.health.belgium.be/sites/default/files/uploads/fields/fpshealth_theme_file/hgr_9339_burnout_zisa4_full_0.pdf

2.1), a distinction was made between core symptoms and secondary, atypical symptoms (Schaufeli & Enzmann, 1989). Inability to perform, as indicated by exhaustion and concomitant cognitive and emotional impairment, constitutes the core of burnout together with the unwillingness to perform, as indicated by mental distance. In fact, cognitive and emotional impairment can be considered as particular aspects of exhaustion because one's energy is lacking for effectively regulating cognitive and emotional processes. Thus the following four dimensions constitute the core of burnout (see figure 2).

- ***Exhaustion.*** This refers to a severe loss of energy that results in feelings of both physical (tiredness, feeling weak) and mental (feeling drained and worn-out) exhaustion. Specific symptoms include; lack of energy to start the new working, feeling completely used-up after a whole day of working, feeling tired quickly even after spending minimal effort at work, and inability to relax after work.
- ***Emotional impairment.*** This manifests itself in intense emotional reactions and feeling overwhelmed by one's emotions. Specific symptoms include; feeling frustrated and angry at work, irritability, overreacting, feeling upset or sad without knowing why, and feeling unable to control one's emotions at work.
- ***Cognitive impairment.*** This is indicated by memory problems, attention and concentration deficits and poor cognitive performance. Specific symptoms include; difficulties to think clearly and learn new things at work, being forgetful and absent-minded, indecision, poor memory, attention and concentration deficits, and trouble staying focused at work.
- ***Mental distance.*** Psychologically distancing oneself from the work is indicated by a strong reluctance or aversion to work. One withdraws mentally – and sometimes even physically – from work and avoids contact with others, for example with customers, clients, and colleagues. Indifference and a cynical attitude are characteristic of mental distance. Little or no enthusiasm and interest for the work exists and one feels that one functions on autopilot.

In addition to these core symptoms, burnout also includes (see figure 2):

- ***Psychological distress.*** This refers to non-physical symptoms that are the result of a psychological problem, such as sleep problems, worrying, feeling tense

and anxious, feeling disturbed by noise and crowds, and weight fluctuations without being on a diet.

- ***Psychosomatic complaints.*** This refers to physical complaints that cannot be explained by a physical disorder, but are exacerbated by or result from some psychological problem. Examples are, palpitations and chest pain, stomach and intestinal problems, headaches, muscle pains and getting sick often.
- ***Depressed mood.*** This refers to a gloomy and sad mood and to the inability to experience pleasure. Depressed people feel powerless, suffer from guilt and are disappointed in themselves. Please note that depressed mood is a normal, temporary reaction to disappointment or loss and should be distinguished from mood disorder, which is a psychiatric syndrome.

Burnout can be described as a work-related[9] condition among employees, who have worked productively and without mayor problems for a long period to the satisfaction of themselves and others. It is defined as: *a work-related state of exhaustion that occurs among employees, which is characterized by extreme tiredness, reduced ability to regulate cognitive and emotional processes, and mental distancing. These four core dimensions are accompanied by depressed mood as well as by non-specific psychological and psychosomatic distress symptoms*.

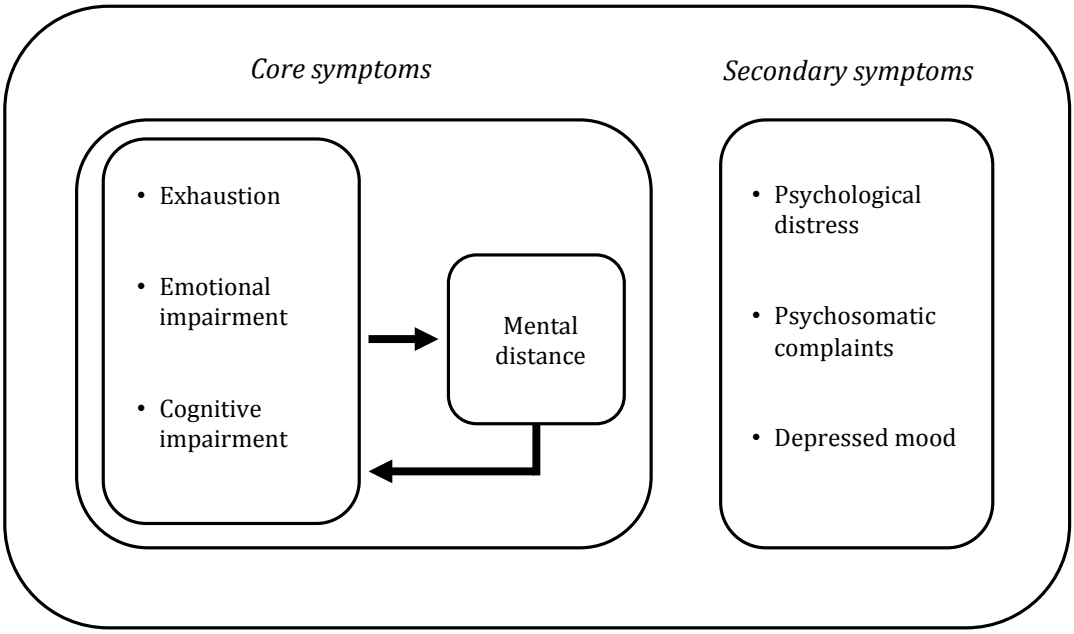
Burnout is caused by an imbalance between high job demands and insufficient resources (Schaufeli & Taris, 2014). Problems outside work and/or personal vulnerability may facilitate the development of burnout. Ultimately, burnout leads to feelings of incompetence and poor work performance (Swider & Zimmerman, 2010).

The intrapersonal dynamic of burnout can be understood as follows (see figure 2). Because one feels extreme tired the necessary energy is lacking for adequately regulating one's emotional and cognitive processes. In other words, the functional capacity for regulating emotional and cognitive processes is impaired. This is subjectively experienced as impairment or loss of emotional and cognitive control. By way of self-protection and in order to prevent further energy depletion and loss of

[9] Psychologically speaking, work refers to any kind of structured, goal-oriented activity that is compulsory in nature and aims to transform the physical or social environment and/or the self. Examples are studying, volunteering, and sports, whilst child rearing, maintaining (partner) relationships, leisure, and housework are excluded.

control, mental distancing occurs. For instance, by developing a detached, indifferent and cynical attitude towards one's work, which, after all, is the root cause of one's exhaustion and emotional and cognitive impairment.

Figure 2: The conceptualization of burnout as used in the BAT



However, this self-protective response is bound to fail because mental distancing evokes negative reactions from others and jeopardizes work motivation and job performance, thereby increasing instead of decreasing stress. So rather than reducing exhaustion and increasing the functional capacity to control one's emotions and cognitions, mental distancing becomes an inherent part of the burnout syndrome itself. Because of feeling exhausted effective (emotional and cognitive) self-control as well as effective control over the work situation (reducing job demands and increasing job resources) is undermined. This is further reinforced by mental distancing, which, in fact, is an inadequate coping strategy. In its turn, the experienced loss of control may further trigger feelings of depression and is accompanied with psychological distress and psychosomatic complaints, which are considered secondary burnout symptoms.

4.4. The items of the BAT

Based on the above description of the core and secondary symptoms of burnout, the authors initially formulated 90 items, 15 per symptom dimension. After elimination of

duplicate and overlapping items, the remaining items were discussed in detail, resulting in the BAT items that are listed below: 23 items refer to core symptoms of burnout (Table 1) and 11 refer to secondary symptoms (Table 2). Because of the pivotal role of the exhaustion dimension, more items were formulated fir this subscale. It was decided *not* to develop a new depression scale because several, short and well-validated scales already exist, such as the 4DKL depression scale, which contains 6 items (Terluin, Van Marwijk et al., 2006).

Table 1: BAT items - core symptoms

Exhaustion

1. At work, I feel mentally exhausted
2. Everything I do at work requires a great deal of effort
3. After a day at work, I find it hard to recover my energy
4. At work, I feel physically exhausted
5. When I get up in the morning, I lack the energy to start a new day at work
6. I want to be active at work, but somehow I am unable to manage
7. When I exert myself at work, I quickly get tired
8. At the end of my working day, I feel mentally exhausted and drained

Mental distance

1. I struggle to find any enthusiasm for my work
2. At work, I do not think much about what I am doing and I function on autopilot
3. I feel a strong aversion towards my job
4. I feel indifferent about my job
5. I'm cynical about what my work means to others

Emotional impairment

1. At work, I feel unable to control my emotions
2. I do not recognize myself in the way I react emotionally at work
3. During my work I become irritable when things don't go my way
4. I get upset or sad at work without knowing why
5. At work I may overreact unintentionally

Cognitive impairment

1. At work, I have trouble staying focused
2. At work I struggle to think clearly
3. I'm forgetful and distracted at work
4. When I'm working, I have trouble concentrating
5. I make mistakes in my work because I have my mind on other things

Table 2: BAT items – secondary symptoms

Psychological distress

1. I suffer from palpitations or chest pain
2. I suffer from stomach and/or intestinal complaints
3. I suffer from headaches
4. I suffer from muscle pain, for example in the neck, shoulder or back
5. I often get sick

Psychosomatic complaints

1. My weight fluctuates without being on a diet
2. I have trouble falling or staying asleep
3. I tend to worry
4. I feel tense and stressed
5. I feel anxious and/or suffer from panic attacks
6. Noise and crowds disturb me

To investigate the best way of scoring the BAT items, three answering scales were compared:

- 5-point relative severity scale; 1 (much less burdensome than normal), 2 (less burdensome than normal), 3 (similar burdensome than normal), 4 (more burdensome than normal), 5 (much more burdensome than normal);
- 7-point frequency scale: 1 (never), 2 (sporadic), 3 (occasional), 4 (often), 5 (very often), 6 (always);

- 5-point frequency scale: 1 (never), 2 (sometimes), 3 (regular), 4 (often), 5 (always).

The first answering scale assumes that burnout symptoms develop progressively over time; initially employees work productively and without problems, but gradually this changes and finally a state of mental exhaustion arises. The second answering scale is used by the MBI version and the third answering scale is more or less generally used by other similar psychological questionnaires.

Three samples of 150 respondents each were drawn by a commercial surveying agency from the Dutch-speaking working population of Belgium. Different BAT answering scales were used in each of these samples. In addition to BAT, the UBOS (Dutch version of the MBI) was also administered, as well as some job stressors (e.g., workload and emotional demands), job resources (e.g., social support and job control), and organizational outcomes (e.g., organization commitment and turnover intention).

A systematic comparison of the three BAT answering formats showed that the 5-point relative severity scale did *not* meet expectations. There were also problems with the 7-point frequency scale, whereas the 5-point frequency scale turned out to be the best answering format. Therefore, this is used in the final version of the BAT.

4.5. The general version of the BAT

Burnout is defined as a work-related mental state, whereby work is not restricted to paid employment but viewed from a broader, psychological perspective (*cf*. footnote 9). This means that work refers to *all* structured, goal-directed activities that are mandatory in nature. For example, psychologically speaking the activities of athletes, volunteers, and students can be seen as "work", and hence, they may also suffer from burnout (Schaufeli, 2018). In contrast, housewives/men, retirees and the unemployed do not, as such, carry out structured, goal-directed and mandatory activities and therefore cannot suffer from burnout. Yet they too may feel exhausted and may suffer from cognitive or emotional impairment. However, that is not because of the specific *activities* they perform, but because of their social *role*, which is characterized by the

lack of a clear focus or goal and a fixed time-structure. In addition the activities they perform are basically voluntary and not mandatory.

A problem arises when measuring burnout among employees who no longer work, for instance, because they are sick listed. Evidently, they cannot answer most BAT items that are displayed in Table 1 because these refer to work. Nevertheless, it is also important to measure exhaustion, mental distance and cognitive and emotional impairment among employees who have dropped out from work due to burnout; for example, when monitoring the progress of a treatment, counseling or return to work program.

This monitoring is possible by using a context-free, general version of the BAT that does not explicitly refer to work (see Table 3). Mental distance regarding work is included in the general version of the BAT as well because its items refer a person's relationship to work rather than to work itself. Therefore, the mental distance items of the BAT can also be answered by those who are currently sick listed. Only the item "At work I don't think much and I function on autopilot" was removed, because it directly refers work covert behavior at work.

The secondary burnout symptoms of the BAT do not refer to work and can therefore also be answered by those who do not work. The same holds for depressed mood (4DKL).

Table 3: The general version of the BAT

Exhaustion

1. I feel mentally exhausted
2. Everything I do requires a great deal of effort
3. I find it hard to recover my energy
4. I feel physically exhausted
5. When I get up in the morning, I lack the energy to start the new day
6. I want to be active, but somehow I am unable to manage
7. When I exert myself , I quickly get tired
8. At the end of the day, I feel mentally exhausted and drained

Mental distance

4. I struggle to find any enthusiasm for my work
5. I feel a strong aversion towards my job
6. I feel indifferent about my job
7. I'm cynical about what my work means to others

Emotional impairment

7. I feel unable to control my emotions
8. I do not recognize myself in the way I react emotionally
9. I become irritable when things don't go my way
10. I get upset or sad without knowing why
11. I may overreact unintentionally

Cognitive impairment

1. I have trouble staying focused
2. I struggle to think clearly
3. I'm forgetful and distracted
4. I have trouble concentrating
5. I make mistakes because I have my mind on other things

5. Conclusions

The most commonly used definition of burnout and the measuring instrument that is based on it are subject to criticism. Questions have raised about the conceptual basis of the Maslach Burnout Inventory (MBI) – known in the Netherlands and Flanders as the Utrecht Burnout Scale (UBOS) – including the nature of its dimensions. There are also a number of technical and psychometric problems with the MBI as well as practical difficulties in interpreting the scores.

These criticisms prompted the development of a new burnout instrument – the Burnout Assessment Tool (BAT) – that can be used for individual and organizational assessment. The starting point for developing the BAT is the theoretical notion that burnout refers to a combination of the inability and unwillingness to perform, represented by its

energetic and motivational dimension, respectively. The former is experienced as exhaustion and the latter as mental distance towards work.

Based on approximately 50 in-depth interviews with professionals who deal with burnout on a daily basis, as well as an inventory of existing burnout questionnaires, four core dimensions of burnout were distinguished. In addition to exhaustion and mental distance, emotional and cognitive impairment emerged as core elements that constitute burnout. In fact, the latter two result from exhaustion because this impedes the effective regulation of emotional and cognitive processes. In other words, in burnout the functional capacity to regulate one's emotional and cognitive processes is affected by a lack of energy.

In addition to these four core symptoms, three secondary symptoms are distinguished that occur not only with burnout, but also with other physical and mental disorders: depressed mood, psychological distress and psychosomatic complaints.

Hence, six short scales were constructed, which together constitute the BAT.

Core symptoms

- Exhaustion (8 items)
- Mental distance (5 items)
- Emotional impairment (5 items)
- Cognitive impairment (5 items)

Secondary symptoms

- Psychosomatic complaints (5 items)
- Psychological distress (6 items)

A separate BAT scale for depressed mood was not developed because other, reliable and valid depression questionnaires exists, such as the 6-item 4DKL depression scale. Although there are also other scales for measuring distress and psychosomatic complaints, it was decided to construct separate BAT scales. Generally speaking, other scales are relatively long and also contain complaints that rarely occur with burnout. For example, the 4DKL distress scale consists of 16 items including: "Did you ever have to do your best to put aside thoughts about any upsetting event(s)?" and

"Did you have to repeat some actions a number of times before you could do something else?".

In addition to a work-related version, a general, context-free version of the BAT is also available that can be filled out by employees who currently do not work but did so until recently. The reason for developing a context-free version is that it makes no sense to ask people about feelings of fatigue at work when they are not working. Yet, they may very well feel exhausted.

Finally, a shortened version has been constructed – both work-related as well as non-work related – each consisting of twelve items, three for each subscale (see Part III).

PART II: PSYCHOMETRIC RESEARCH

1. Validation samples

For the psychometric examination with the BAT, mainly two samples were used (see Table 4):

- *A representative sample of the Flemish working population*. In the summer of 2017, a commercial surveying agency (iVox) drew a random sample from the Flemish workforce that is representative of age, gender and industry. The selection criteria for age and gender were "hard", meaning that they must perfectly match the distribution of the Flemish workforce. Industry, on the other hand, was a "soft" selection criterion, meaning that a maximum deviation of 10% from the Flemish workforce was allowed.
- *A representative sample of the Dutch working population*. Simultaneously with the Flemish sample, also a Dutch representative sample was drawn according a similar procedure and using the same specifications.

In addition, supplementary samples were used for certain specific analyzes. These are described in greater detail in the text below.

Table 4: overview of the representative Flemish and Dutch validation samples

	Flanders (N=1.500)	**Netherlands (N=1.500)**
Gender		
Male	54.3%	51.4%
Female	45.7%	45.9%
Age		
Mean	41.3 yrs	41.3 yrs
SD	11.5 yrs	13.3 yrs
Education		
Primary	.08%	.08%
Intermediary	70,9%	53,8%
Higher	28.3%	45.4%

	Flanders (N=1.500)	Netherlands (N=1.500)
Industry sector		
Primary sector (e.g. agriculture, mining, fishing)	.7%	3.6%
Industrial sector (e.g. construction, chemical industry)	19.9%	18.5%
Service sector (e.g. trade, banking, hospitality)	33.1%	47.0%
Public administration	18.9%	9.5%
Education	12.7%	7.2%
Health & welfare	14.7%	14.3%
Occupational level		
Unskilled (e.g. machine operator, production worker)	2.9%	8.8%
Skilled (e.g. electrician, welder)	9.9%	14.5%
Administrative and operative (e.g. shop assistant, secretary)	22.9%	22.7%
Lower middle management (e.g. teacher, sales manager)	38.1%	15.1%
Higher middle management (e.g. physician, lawyer)	22.3%	27.5%
Higher management (e.g. school principle, CEO)	6.9%	11.3%
Type of job		
Fulltime	78.3%	64.1%
Parttime	21.7%	35.9%
Work hours/week – according to labor contract		
Mean	34.5	30.4
SD	8.3	12.0
Work hours/week – actual state		
Mean	39.4	33.5
SD	9.5	13.1
Job tenure		
Mean	18.7 yrs	18.3 yrs
SD	11.7 yrs	12.5 yrs
Treated for burnout in the past 5 years?		
Yes	6.5%	8.7%
No	93.5%	91.3%

The representativeness of the two samples is guaranteed by the way of sampling that was used, at least as far as age, gender and industry are concerned. In terms of gender, age and work experience, both samples are virtually the same. However, the Dutch sample is higher educated. Respondents in the Flemish sample are more often employed in public administration and education, while Dutch respondents work more often in the service sector. The Dutch sample has both a lower (skilled and

unskilled) and a higher (higher middle management and higher management) occupational level, while Flanders is overrepresented in lower middle management. Part-time work is more often performed in the Netherlands, so that the average number of hours worked per week – both according to the labor contract as well in as in reality – is lower than in Flanders. Finally, compared to the Flemish respondents, the Dutch respondents are somewhat more often treated for burnout in the past 5 years.

2. Score distribution of the items

First of all, it was checked to what extent the scores on the items of the BAT are normally distributed. Table 5 provides an overview of the most important distribution characteristics.

Table 5: Range of the mean (M), standard deviation (SD), skewness and kurtosis of the BAT items for Flanders (N = 1,500) and the Netherlands (N = 1,500)

	M	SD	Skewness	Kurtosis
Core symptoms (23 items)				
Flanders	1.66 – 2.31	.74 – 1.04	.29 – 1.16	-.49 – .88
Netherlands	1.91 – 2.60	.95 –1.13	.17 – .91	-.85 – .26
Secondary symptoms (11 items)				
Flanders	1.74 – 2.80	.79 – 1.14	.07 – 1.21	-.91 – .91
Netherlands	1.86 – 2.41	1.00 – 1.18	.37 – 1.01	-.93 – .22

At first glance, the distribution characteristics do not give cause for concern. Because of the large sample sizes it does not make sense to conduct a formal test for normality, nor to use critical values for skewness and kurtosis (< 1.96, after z-score transformation). With a sample size of 1,500, even very small and irrelevant deviations from normality will be statistically significant (Field, 2013; pp. 183-186). In case of large samples, normal distribution can best be assessed by visual inspection.

It appears that the items that tap exhaustion and cognitive impairment are fairly normally distributed, with the score of 2 occurring the most and the scores of 4 and 5 the least frequent. The score distribution of the items that refer to mental distance and emotional impairment, on the other hand, show a different picture with slight deviations from normality; the lowest score (1) occurs most frequently and the score frequency decreases successively as the scores increase. The score distribution of the secondary burnout symptoms shows a mixed picture; most items are fairly normally

distributed, but sometimes also a different pattern is found, comparable with mental distance and emotional impairment. The patterns of score distributions for the core and secondary symptoms are roughly the same for the Netherlands and Flanders.

It can be concluded that the scores for exhaustion and cognitive impairment are approximately normally distributed. However, this is not the case for the other two core symptoms; mental distance and emotional impairment. Respondents generally do not exhibit these burnout symptoms. Some secondary burnout symptoms are normally distributed, while others are not.

By the way, the statistical analyzes used in this manual, such as correlations, regressions, t-tests and analyses of variance, are fairly robust for violations of normality. In other words, the score distribution of the BAT-items does not preclude using these types of analysis.

3. Factorial validity

Because different aspects of burnout can be distinguished, a number of confirmatory factor analyzes have been carried out. A distinction was made between core and secondary symptoms, which were initially analyzed separately.

3.1. Core symptoms of burnout

Three models including core symptoms of the BAT were tested:

1. A 1-factor model in which all 23 items are supposed to load on one underlying factor (Model 1);
2. A 4-factor correlated model, in which the four core aspects (exhaustion, mental distance, and cognitive and emotional impairment) are distinguished but correlated with each other (Model 2);
3. A second-order model, assuming that the four core aspects refer to one underlying factor (Model 3). This model corresponds to the view that burnout is a syndrome that consists of four types of complaints, which refer to one underlying psychological state or disorder.

In order to determine to what extent the three models fit the data from both samples, they were tested simultaneously in the Flemish and Dutch sample, using the so-called multiple group method.

Table 6 shows the fit indicators of the models. To assess the goodness-of-fit, it is recommended to use different indicators (Byrne, 2010). A good fit exists when the Comparative Fit Index (CFI) and the Tucker-Lewis Index are at least higher than .90 – and preferably higher than .95 – and the Root Mean Square Error of Approximation (RMSEA) is .08 or less (Hu & Bentler, 1995; pp. 76–99).

Table 6: Confirmatory factor analysis of the core dimensions of the BAT (Flanders, N = 1,500 and the Netherlands, N = 1,500)

Model		χ^2	df	CFI	TLI	RMSEA
1	***1-factor model***	13725.42	460	.78	.76	.10
2	***4-factor model***	2471.28	448	.97	.96	.04
3	***Second-order model***	2556.65	452	.97	.96	.04

Note: χ^2 = chi-square; df = degrees of freedom; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation.

Models 2 and 3 both fit well with the data, which means that the four aspects of burnout can indeed be distinguished. All items load high on the four factors, ranging from .76 to .88 in Flanders and from .74 to .91 in the Netherlands. However, these four aspects are closely related, as is shown in Table 7. In the Dutch sample, inter-correlations between the factors are at least .75 and in the Flemish sample at least .60.

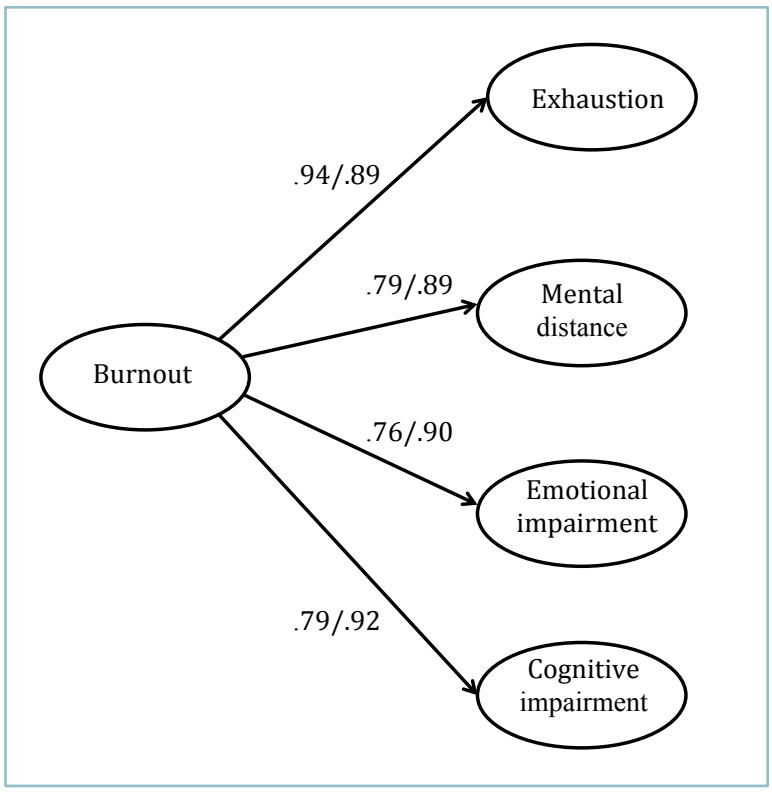
Table 7: Correlations between the latent BAT factors in Flanders (N = 1,500) and the Netherlands (N = 1,500)

Scale		1	2	3	4
1	***Exhaustion***	--	.83	.77	.81
2	***Mental distance***	.74	--	.78	.79
3	***Emotional impairment***	.71	.61	--	.85
4	***Cognitive impairment***	.74	.60	.61	--

Note. Values for the Netherlands above the diagonal and for Flanders below the diagonal; all correlations are significant (p <.001).

This strong interrelation is also reflected in the second-order model (Model 3). Figure 3 shows the extent to which the four aspects refer to the same underlying factor that can be interpreted as a burnout.

Figure 3: The structure of the BAT



Note: Flanders/Netherlands

In practical terms, this means that the BAT can be understood both as a one-dimensional questionnaire that measures burnout on the basis of a single total score, *and* as a four-dimensional questionnaire that measures four strongly interrelated aspects of burnout. Therefore, in the rest of this manual, the total score of the BAT is used as well as the scores on its four subscales, which can be interpreted as specific aspects or dimensions of burnout.

3.2. Secondary burnout symptoms

Two models were tested concerning the secondary symptoms: (1) a 1-factor model, in which all items load on one single factor (Model 1); (2) a correlated 2-factor model in which psychological distress and psychosomatic complaints each load on a separate factor (Model 2). A second-order model with only two factors is mathematically not identified and can therefore not be tested. Again, the models were tested simultaneously in Flanders and the Netherlands (Table 8).

Table 8: Confirmatory factor analysis of the secondary burnout symptoms of the BAT (Flanders, N = 1,500 and the Netherlands, N = 1,500)

Model		χ^2	df	CFI	TLI	RMSEA
1	*1-factor model*	1625.10	88	.91	.89	.08
2	*2-factor model*	939.60	86	.95	.94	.06

Note: χ^2 = chi-square; df = degrees of freedom; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation.

Model 2 fits the data somewhat better than Model 1, although two of the three fit indicators of Model 1 meet the criteria for good fit. Both factors are highly correlated; .81 in Flanders and .90 in the Netherlands. It is also striking that the item "My weight fluctuates without being on a diet" loads in the Flemish sample poorly on the psychosomatic complaints factor (.40). The remaining items load between .62 and .78 on both factors in Flanders and between .72 and .81 in the Netherlands.

Because the depression subscale of the 4DKL was administered in the Flemish sample only, factor models including depression could only be tested in this sample. Table 9 shows the fit of a 1-factor model on which all secondary burnout items are assumed to load, including depressed mood (Model 1), as well as a 3-factor model with psychological distress, psychosomatic complaints and depressed mood as separate factors (Model 2).

Table 9: Confirmatory factor analysis of the secondary burnout symptoms of the BAT and the 4DKL depression scale (Flanders, N = 1,500)

Model		χ^2	df	CFI	TLI	RMSEA
1	*1-factor model*	5507.05	119	.68	.64	.17
2	*3-factor model*	1695.39	116	.90	.88	.09
2a	*3-factor model (adjusted)*	812.14	115	.96	.95	.06
3	*2-factor model*	1209.79	117	.93	.92	.08

Note: χ^2 = chi-square; df = degrees of freedom; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation.

The 1-factor model (M1) clearly does *not* fit the data, but the fit of the 3-factor model (M2) also does not meet all criteria for good fit. Inspection of the so-called Modification Indices reveals that the measurement errors of two depression items are correlated. The content of both items strongly overlaps because both refer to death ("If only I were dead" and "That it would be better if you were only dead"). If the correlation between the measurement errors of these two items is assumed in a modified model (M2a), the fit improves considerably and all criteria are met. All items load high on their

corresponding factors in Model 2a – ranging from .61 to .87 – except "My weight fluctuates without dieting," which only loads .40 on distress. Moreover, the two BAT scales correlate very high with each other (*r* = .81), but far less high with the depression scale of the 4DKL; *r* = .64 and *r* = .54 for distress and psychosomatic complaints, respectively. This agrees with the poorer fit of the 2-factor model (M3) in which the two BAT scales are combined. Although the fit indices of this model are somewhat less good than those of the adjusted 3-factor model (M2a), all fit indices meet the criteria. Both factors of Model 3 correlate moderately (*r* = .64) with each other.

It can be concluded that the two types of secondary symptoms of the BAT are so closely related that, practically speaking, it makes little sense to distinguish them. This is also confirmed by an exploratory factor analysis, which yields only one factor in both Flanders and the Netherlands that explains 49% and 61% of the variance in both countries, respectively. In addition, all items load higher than .65 on this single factor (again with the exception of "My weight fluctuates without dieting"). Apparently, psychological distress and psychosomatic complaints are closely intertwined, which is not the case for depressed mood. Although depressed mood is positively related with distress and psychosomatic complaints, the association is not very strong. Therefore, in the remainder of this manual, one scale for secondary psychological and psychosomatic symptoms or distress will be used.

3.3. Relationship between core and secondary burnout symptoms

Table 10 shows the correlations between the scales that measure the core and secondary burnout symptoms. As mentioned above, the depression scale of the 4DKL was only administered in Flanders. The observed correlations between the scales, as displayed in Table 10, are by definition lower than those between the latent factors in Table 6 because the latter are uncontaminated correlations without measurement errors. Unsurprisingly, the four core symptoms of burnout are strongly related with the total score on the BAT (.80 < r <.93), but less strongly with the total score of the secondary symptoms (.50 < r <.75). Moreover, both total scores are strongly interrelated: in Flanders (r = .71) and in the Netherlands (r = .80).

Table 10: Correlations between the core and secondary burnout symptoms in Flanders (N = 1,500) and the Netherlands (N = 1,500)

Scale		1	2	3	4	5	6
1	***Exhaustion***	--	.78	.73	.76	.93	.75
2	***Mental distance***	.66	--	.73	.74	.89	.65
3	***Emotional impairment***	.65	.55	--	.81	.89	.73
4	***Cognitive impairment***	.68	.55	.59	--	.90	.74
5	***Total core symptoms***	.92	.82	.80	.81	--	.80
6	***Total secondary symptoms***	.71	.50	.60	.55	.71	--
7	***Depressed mood***	.56	.52	.52	.44	.61	.56

Note: Values for the Netherlands above the diagonal and for Flanders below the diagonal; all correlations are significant ($p < .001$).

This means that the core and secondary symptoms overlap 50% in Flanders and 64% in the Netherlands. Depressed mood is roughly as closely related to core ($r = .61$) as to secondary burnout complaints ($r = .56$).

An explorative factor analysis in the Flemish sample including the four scale scores of the core symptoms, the total score on the secondary symptoms and the 4DKL depression scale yields one, common factor. This explains 65% of the variance; exhaustion loads highest (.89) and depressed mood lowest (.74) on the single factor. The same analysis with the Dutch sample, albeit without the depression scale, similarly yields one, common factor that explains 79% of the variance. All scales load higher than .75 on this single factor and again exhaustion loads highest (.81). Taken together, these results correspond to the notion that burnout is a syndrome with core and secondary symptoms, of which exhaustion is the principal element.

4. Classification of BAT scales

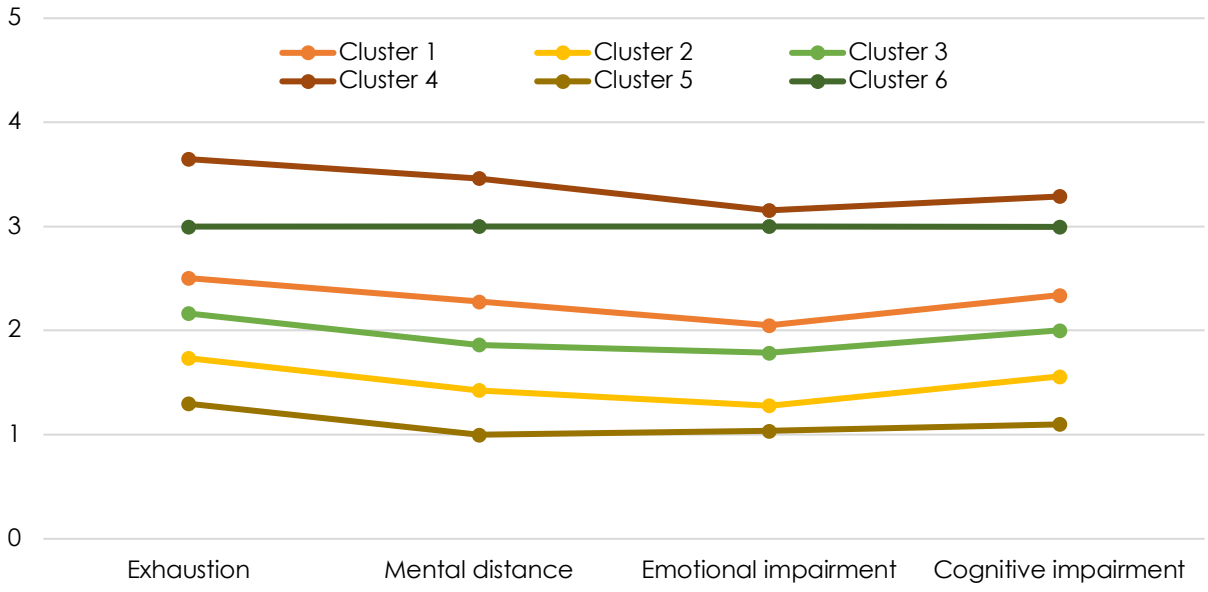
To determine whether it is appropriate to use a single BAT score, a Latent Class Analysis (LCA; Hagenaars & McCutcheon, 2009) was performed using the core dimensions of the BAT. The LCA model that was tested assumed that the BAT consists of four latent factors (exhaustion, mental distance, and emotional and cognitive impairment), which, in their turn, are each composed of a number of items. So the subscales of the BAT are represented in the LCA by latent variables. The fit of this latent 4-factor model is calculated as a function of the number of classes or clusters.

The LCA was performed simultaneously on all 3,000 respondents from Flanders and the Netherlands, using country as a covariate. It appears that the model with 6 clusters fits

best with the data: (BIC (LL) = 15645.25; AIC (LL) = 15326.91; SABIC (LL) = 15476.85; classification error = .07.

As can be seen from Figure 4, these 6 clusters differ in *level* and not in pattern of BAT subscale scores. This means that respondents from each of the clusters differ in the extent to which they report burnout complaints, varying from very high (Cluster 4) to very low (Cluster 5). Remarkably, no pattern is observed in which respondents from a certain cluster score high on some particular BAT scale(s) and low some other. The fact that the observed clusters only differ regarding to the *level* of the scores and not with regard to the score *pattern*, constitutes an additional argument for the use of the total score of the BAT.

Figure 4: Average BAT scores on the 6 clusters



Note. Cluster 1 = "High average"; Cluster 2 = "Low"; Cluster 3 = "Low average"; Cluster 4 = "Very high"; Cluster 5 = "Very low", Cluster 6 = "High".

Of course, the mean scores on the BAT scales differ significantly between groups (exhaustion: $F_{(2994, 5)} = 1278.86$, $p < .001$; mental distance $F_{(2994, 5)} = 1124.04$, $p < .001$; emotional impairment, $F_{(2994, 5)} = 1053.24$, $p < .001$; cognitive impairment, $F_{(2994, 5)} = 1195.10$, $p < .001$). Post-hoc tests on the scales also show significant differences between each of the clusters ($p <. 001$).

But do the clusters also differ with regard to other variables? Subsequent analysis of variance shows that the 6 clusters differ systematically in terms of work engagement ($F_{(2994, 5)}$ = 136.70, p < .001), age ($F_{(2994, 5)}$ = 39.65, p < .001), job tenure ($F_{(2994, 5)}$ = 24.88, p <.001), secondary burnout complaints ($F_{(2994, 5)}$ = 482.08, p < .001), and number of days on sick leave ($F_{(2994, 5)}$ = 4.63, p < .001) . The pattern is always the same; as burnout increases work engagement, age and job tenure decrease, and secondary symptoms and sickness absence increase.

Furthermore, the differences between the clusters show a monotonous trend; that is, the decrease or increase is linear in nature, running from the highest scoring cluster (# 4) to the lowest scoring cluster (# 5). Furthermore, it appears that 46% in the (very) high versus 16% in the (very) low group is currently on sick leave ($\chi^2_{(5)}$ = 94.32, p < .001) and that 48% against 11% have been treated for burnout in the last 5 years ($\chi^2_{(5)}$ = 172.64, p <.001).

We also looked at differences in job demands and job resources between the 6 clusters[10]. To this end, two MANOVAs were carried out in the Dutch sample, both of which showed a significant effect of job demands ($F_{(1483, 12)}$ = 5265.57, p < .001) and job resources ($F_{(1472, 23)}$ = 1611.17, p < .001). It appears that scores on most job demands increase and on most job sources decrease monotonously as the level of burnout rises. This applies to emotional and physical demands, bureaucracy, role conflict and harassment (all job demands), as well as to the use of skills, availability of tools, value conflict, role clarity, support from colleagues and supervisor, meeting expectations, team atmosphere, team effectiveness, task variety, job control, person-job fit, learning and development opportunities, organizational justice and trust in management (all job sources). There is also more work-life conflict at higher burnout levels. Typically, the difference between the "very high" and "high" groups is not significant for most job resources.

Finally, we analyzed differences in well-being and personality between the 6 clusters[11]. To this end, two MANOVAs were performed in the Flemish sample, which both showed

[10] For a description of job demands and job resources; see Section 7 below and Appendix 1.
[11] For a description of well-being and personality; see Sections 6 and 7 below respectively.

a significant effect for well-being ($F_{(1490, 5)} = 1276.15$, $p < .001$) and personality ($F_{(1490, 5)} = 28215$, $p < .001$). It appears that the scores on the various aspects of well-being are less favorable as the level of burnout rises. The higher the cluster scores on burnout, the higher the scores on work addiction (working excessively and compulsively), boredom and depression (as measured with the BDI; Beck, Steer & Brown,1996), and the lower the score on job satisfaction. It also appears that the degree of extraversion, agreeableness and conscientiousness decreases as the level of burnout increases, while the degree of neuroticism increases. Again, it is noteworthy that the differences between the two highest scoring burnout clusters often lack significance; this applies to boredom, depression, extraversion, agreeableness and conscientiousness.

In summary, it seems that indeed a single BAT score can be used. After all, 6 groups or clusters were found with similar scoring levels relative to each other on each of the BAT subscales, ranging from very low to very high. These groups systematically differ from each other; employees in groups with higher burnout levels experience higher job demands and less job resources, and feel less well and healthy. Moreover, in groups with higher burnout levels relatively many employees are found who are currently on sick leave and have been treated for burnout. All in all, these results support the validity of the empirical grouping in 6 clusters that differ in levels of burnout.

5. Reliability

Two types of reliability were examined: internal consistency (Cronbach's α) and test-retest reliability (stability coefficient r_t). In addition, it was investigated to what extent the individual's own score on the BAT agrees with that of someone else; the inter-rater reliability.

5.1. Internal consistency

Table 11 shows the values of Cronbach's α, which indicate the extent to which the BAT scale is internally consistent. In general, α-values $\geq .70$ are considered sufficient and values $\geq .80$ good (Nunally & Bernstein, 1994). All scales have an excellent internal consistency. With one exception, all items contribute to the internal consistency of the

scale to which they belong. That means that the internal consistency does not increase when one or more items are removed from any scale.

Table 11: The internal consistency of the BAT scales (coefficient α)

Scale	# items	Flanders (N = 1,500)	Netherlands (N = 1,500)
Exhaustion	8	.92	.94
Mental distance	5	.91	.93
Emotional impairment	5	.90	.94
Cognitive impairment	5	.92	.94
Total core symptoms	23	.96	.97
*Total secondary symptoms**	11	.89	.94
Depressed mood	6	.93	--

Note: * In the Flemish sample, the 10-item version has an a value of .90.

The only exception is the item "My weight fluctuates without being on a diet." If this item is deleted from the secondary symptom scale, the value of α increases slightly from .89 to .90 in Flanders, while it remains unchanged in the Netherlands. This is in accordance with the previous section, which showed that this item had a relatively low factor loading, particularly in Flanders. It was therefore decided *not* to include this item in the scale that measures secondary burnout symptoms.

5.2. Test-retest reliability

In a group of 964 Flemish respondents who were approached online[12] (49.7% female, 50.3% male, age = 40.3, SD = 10.5), the BAT was administered three times, with an interval of six months. At T2 and T3, 649 and 481 respondents answered the questionnaire, respectively. This means that the dropout between T1 and T2 was 32.6% and between T2 and T3 was 25.8%. The dropout over the entire one-year period was 50%.

Table 12 shows that stability coefficients of the BAT subscales range from .60 to 75 over a 6-month period, and from .54 to .69 over a 1-year period. That makes sense because as time goes by, stability decrease of course

[12] Using the website http://burn-out.vlaanderen

Table 12: Test-retest reliability of the BAT scales (stability coefficient r_t)

Scale	T1 -T2 (N = 597)	T2-T3 (N = 368)	T1-T3 (N=447)
Exhaustion	.71	.75	.69
Mental distance	.68	.64	.60
Emotional impairment	.67	.64	.60
Cognitive impairment	.62	.66	.54
Total core symptoms	.74	.73	.68
Total secondary symptoms	.80	.82	.80
Depressed mood	.66	.70	.64

Furthermore, it appears that: (1) the test-retest reliability of the total BAT (core symptoms) is higher than that of its subscales; (2) the stability of the secondary symptoms is highest and; (3) the stability of depressive feelings (4DKL) is comparable to that of the BAT subscales.

There is no criterion for sufficient test-retest reliability because this depends on the nature of the construct and the purpose of the measurement. The values of Table 12 indicate on the one hand that the scores on the BAT are reasonably stable across a period of 6 months, which corresponds with the chronic nature of burnout symptoms. On the other hand, however, there is also a certain variation in time, which indicates that the BAT is sensitive to change.

In short, the stability coefficients of the BAT's core symptoms are in the mid-range; between 40% and 55% of the variance remains stable over a half-year period, and between 35% and 45% over the period of one year. This also applies to the stability of depressive feelings, as measured by the 4DKL. The stability of the secondary symptoms is highest with over 60% explained variance over both the short and longer period.

5.3. Inter-rater reliability

For 23 employees (39% men and 61% women; mean age 38.3 years, SD = 11.6) working in mental health care, their partner was also asked to complete the BAT with the instruction that the questions do not concern themselves but their working partner (De Bie, 2018). On the basis of this peer rating, an independent judgment was obtained about the degree of burnout of the focal employee in the eyes of his or her partner.

Table 13 below shows the agreement in scoring of the focal person and his or her partner in terms of correlation coefficients (*r*) and Cohen's *d*'s, an association measure that is independent of sample size. The internal consistencies (coefficient α) of the BAT scales are very good and vary from .82 to .95 among the focal employees and from .87 to .96 among their partners.

Table 13: Inter-rater reliability of the BAT (N = 23)

Scale	*r*	*d*
Exhaustion	.63**	.37
Mental distance	.69***	.19
Emotional impairment	.60**	.23
Cognitive impairment	.12	.18
Total core symptoms	.63**	.19
Total secondary symptoms	.70***	.14

Note: ** p <.01, *** p < .001.

In general, the inter-rater reliability is low to moderate. Perhaps this is due to the small number of respondents. Particularly the very low agreement for cognitive impairment is striking. Perhaps cognitive impairment occurs in particular at work and not so much at home where it could have been observed by the partner. Alternatively, indications of cognitive impairment such as lack of attention, poor concentration, and not being able to think clearly are covert and hence may thus, because of their very nature, not be observed by others.

By the way, the inter-rater reliability of the BAT appears to be better than that of the MBI, with the exception of cognitive impairment. For example, correlations between MBI self-reports and peer ratings from family members range from .20 to only .56 in different studies among social workers (N = 91), physicians (N = 43), nurses (N = 180), and police officers (N = 142) (Maslach & Jackson, 1986).

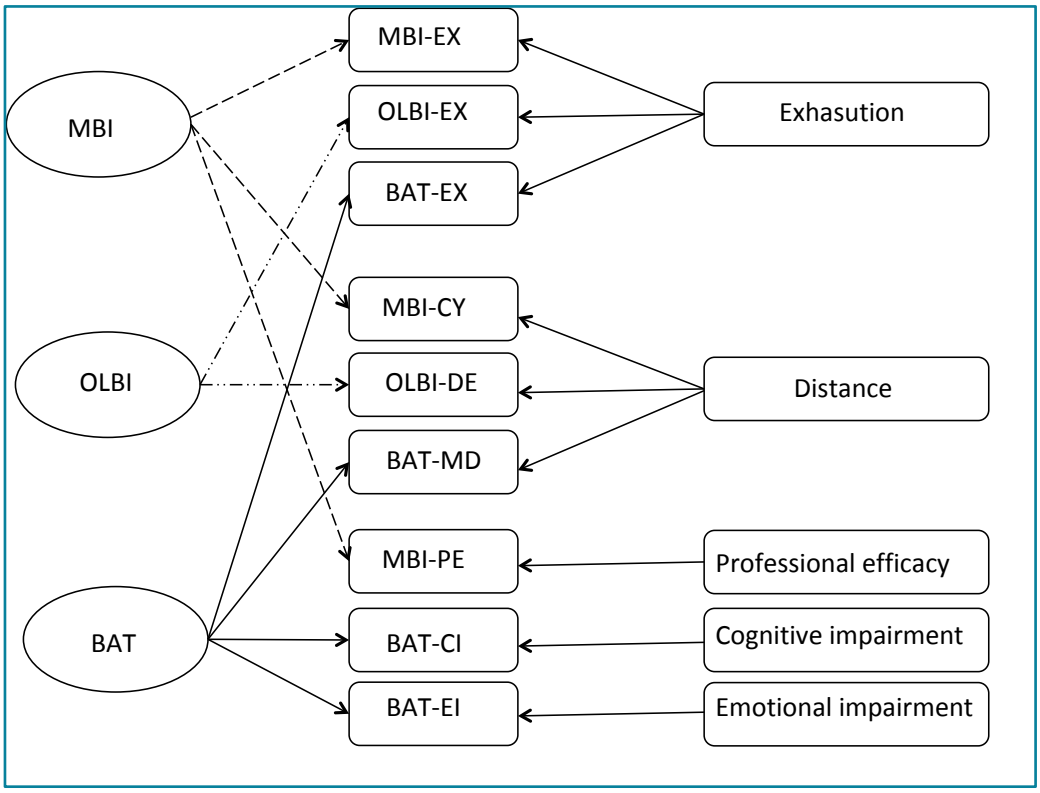
6. Convergent and discriminant validity

It was investigated to what extent the BAT measures the same construct as other commonly used burnout scales (convergent validity) and to what extent scores on the BAT can be distinguished from scores on questionnaires that measure other concepts, whether or not burnout-related (discriminant validity).

6.1. Other burnout questionnaires

It is important to examine the relationship between the BAT and other, commonly used burnout questionnaires such as the UBOS (Schaufeli & Van Dierendonck, 2000) and the OLBI (Demerouti et al., 2003). In order to determine the convergent and discriminant validity of the BAT vis-à-vis these questionnaires, the Multi-Trait, Multi-Method approach of Widaman (1985) was used. According to this approach the *Correlated Trait-Correlated Methods Model* (CT-CM) – Model 1 – constitutes the basic model with which all other models are compared (see Figure 5).

Figure 5: The Multi-Trait, Multi-Method framework for the BAT, UBOS and OLBI (Flanders N = 1,500)



Note: EX = exhaustion, CY = cynicism, DE = disengagement, MD = mental distance, CI = cognitive impairment, EI = emotional impairment.

Model 2, the *Correlated Traits -No Method Model* (CT-NM) assumes that the properties (traits) that each instrument pretends to measure best represent the structure of the data. Convergent validity exists when Model 1 fits the data better than Model 2.

Model 3 and Model 4 are variations of Model 1. Model 3, the *Perfectly Correlated Traits- Correlated Method Mode*l (PCT-CM), assumes that the properties are perfectly

correlated, whereas Model 4, the *Correlated Traits-Perfectly Correlated Method Model* (CT-PCM) assumes that the measurement methods (questionnaires) are perfectly correlated.

Both models are used to test the discriminant validity based on the assumption that when the properties (traits) and measurement methods (questionnaires) are independent of each other, a model that allows both to correlate yields a better fit to the data compared to the basic model (Model 1). By comparing Models 3 and 4 with Model 1, discriminant validity with regard to properties and measurement methods are demonstrated, respectively.

Table 14 shows the results of the analyzes with the Flemish sample, in which the OLBI was administered in addition to the BAT and the UBOS. In the Dutch sample only the BAT and the UBOS were administered (see Tables 16 and 17).

Table 14: Fit indicators of the Multi-Trait, Multi-Method framework of the BAT (Flanders; N = 1,500)

Model		χ^2	df	CFI	TLI	RMSEA
1	***CT-CM model***	5803.79	1310	.91	.90	.05
2	***CT-NM model***	10020.64	1367	.83	.82	.07
3	***PCT-CM model***	7022.20	1313	.89	.88	.05
4	***CT-PCM model***	11203.54	1320	.80	.79	.07

Note: χ^2 = chi-square; df = degrees of freedom; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation.

As expected, Model 1 fits the data best and fits significantly better to the data than the other three models. The fact that Model 1 fits better than Model 2 confirms the convergent validity of the BAT with the UBOS and the OLBI. This means that the BAT overlaps with both other burnout instruments.

Table 15: Comparison of the fit of different Multi-Trait, Multi-Method models (Flanders; N = 1,500)

Model comparison	$\Delta\chi^2$	Δdf	p
Convergent validity			
Model 1 vs. Model 2	4134.88	57	< .0001
Discriminant validity			
Model 1 vs. Model 3 (traits)	896.35	3	< .0001
Model 1 vs. Model 4 (methods)	4871.58	10	< .0001

Note.: $\Delta\chi^2$ = chi-square difference; Δdf = difference in degree of freedom.

Furthermore, the fact that Model 1 fits better to the data than Models 3 and 4 confirms the discriminant validity of the BAT with respect to properties and measurement method, respectively. The different properties (traits) are not perfectly correlated and therefore differ from each other, while each of the three questionnaires provides unique information about burnout.

In summary; on the one hand convergence is demonstrated of the BAT with both other burnout questionnaires regarding exhaustion and mental distance, but on the other hand there is also sufficient divergence so that each instrument contributes independently to the measurement of burnout.

The same analyzes were also performed on the data from the Dutch sample, in which only the BAT and the UBOS were administered. For the results, see Tables 16 and 17.

Table 16: Fit indicators of the Multi-Trait, Multi-Method framework of the BAT (the Netherlands; N = 1,500)

Model		χ^2	df	CFI	TLI	RMSEA
1	***CT-CM model***	1673.63	616	.97	.97	.03
2	***CT-NM model***	6364.67	772	.86	.85	.07
3	***PCT-CM model***	1806.56	617	.97	.96	.04
4	***CT-PCM model***	5250.98	626	.87	.86	.07

Note: χ^2 = chi -square; df = degrees of freedom; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation.

Again Model 1 fits best to the data. The finding that Model 1 fits better with the data than Model 2 confirms the convergent validity of the BAT with respect to the UBOS (see Table 17). Moreover, the finding that Model 1 fits better than Models 3 and 4 demonstrates the discriminant validity of the BAT with regard to properties and measurement method, respectively.

Table 17: Comparison of the fit of different Multi-Trait, Multi-Method models (the Netherlands)

Model comparison	$\Delta\chi^2$	Δdf	p
Convergent validity			
Model 1 vs. Model 2	5028.70	156	< .0001
Discriminant validity			
Model 1 vs. Model 3 (traits)	21.64	1	< .0001
Model 1 vs. Model 4 (methods)	2087.75	10	< .0001

Note.: $\Delta\chi^2$ = Chi-square difference; Δdf = difference in degree of freedom.

Taken together, also in the Netherlands convergent validity exists of the BAT with the UBOS regarding exhaustion and mental distance. And like in Flanders, at the same time the BAT also shows sufficient divergence, so that each burnout instrument contributes independently to the measurement of burnout.

6.2. Work engagement, work addiction and boredom

It is not only important to establish the validity of the BAT with respect to other burnout questionnaires, but also the discriminant validity with respect to questionnaires that measure other aspects of work-related well-being, such as work engagement, work addiction, and boredom at work. After all, the BAT is supposed to measure something different than (lack of) work engagement, work addiction or boredom at work.

Work engagement is defined as: "... as a positive, fulfilling, work-related state of mind that is characterized vigor, dedication, and absorption" (Schaufeli, Salanova, González-Romá & Bakker, 2002, p. 74). More specifically; "Vigor is characterized by high levels of energy and mental resilience while working, the willingness to invest effort in one's work, and persistence even in the face of difficulties. Dedication is characterized by a sense of significance, enthusiasm, inspiration, pride, and challenge. The final dimension of engagement, absorption, is characterized by being fully concentrated and deeply engrossed in one's work, whereby time passes quickly and one has difficulties with detaching oneself from work" (pp. 74-75).

Work engagement is seen as the positive counterpart to burnout (Maslach, Schaufeli & Leiter, 2001; Schaufeli & De Witte, 2017). Indeed González-Roma, Schaufeli, Bakker and Lloret (2005) found that emotional exhaustion and cynicism – as measured by the MBI-GS –, and vigor and exhaustion – as measured by the Utrecht Work Engagement Scale (UWES; Schaufeli & Bakker, 2004a) – are the opposite poles of two dimensions: energy and identification. This means that burnout is characterized by a low level of energy (exhaustion) and poor identification with one's work (cynicism), whereas engagement is characterized by a high level of energy (vigor) and strong identification (dedication). In the study by González et al. (2005) reduced professional efficacy was *not* included because it is not considered to be a core element of burnout. Seen from this perspective, it is not surprising that Schaufeli, Taris and Van Rhenen (2008) found that instead of loading on burnout, personal efficacy loads on

work engagement. This illustrates that rather than being a constituent element of burnout, professional efficacy is positively related to work engagement. In summary, it can be expected that scores on the BAT are negatively related to those on the UWES.

Work addiction – also known as workaholism – is the uncontrollable need to work excessively hard (Schaufeli, Van Wijhe, Peeters & Taris, 2011). Work is an obsession for work addicts; they feel useless, guilty, irritated and tense when they are unable to work, but once at work these feelings of discomfort disappear. Work addiction includes a behavioral component (excessive working) as well as a cognitive component (compulsive working). In fact, workaholism is the combination of working excessively and compulsively. Both aspects are measured by the Dutch Workaholism Scale (DUWAS; Schaufeli et al., 2011). Workaholism is expected to be positively associated with burnout because work addicts are, by their very nature, unable to unwind and relax after work, so that they no not sufficiently recover from their efforts. Consequently, they are likely to deplete their energy resources and run the risk of burning out (Maslach, 1986). This is illustrated by a meta-analysis of 18 studies (N = 10,319) that showed a modest, positive association (ρ = .40) between workaholism and burnout (Clark, Michel, Zhdanova & Baltes, 2016).

Boredom at work is a negative psychological state characterized by a low degree of activation (arousal) and a high degree of dissatisfaction (Reijseger, Schaufeli, Peeters, Taris, Van Beek & Ouweneel, 2012). Boredom occurs in low-activating and unchallenging working environments, for example, when monotonous and repetitive tasks have to be completed or when the job requirements are far below the employee's levels of skill and qualification. In contrast to burnout, which results from *over*-stimulation, boredom results from *under*-stimulation. Yet, both are characterized by deactivation, which is the reason why burnout and boredom are positively related to each other. Moreover, it appeared from a study that used the Dutch Boredom Scale (DUBS; Reijseger et al., 2012) that this positive relationship is stronger for cynicism than for exhaustion.

To investigate the convergent and discriminant validity of the BAT in relation to work engagement (UWES), work addiction (DUWAS) and boredom at work (DUBS), the guidelines of Formel and Larcker (1981) were followed. They argue that discriminant

validity exists when the average explained variance (AVE) of a latent factor – in our case burnout – exceeds the squared correlation (R^2) of the target latent factor – in our case work engagement, work addiction and boredom at work. Tables 18 and 19 display the values of the AVE and R^2 for Flanders and the Netherlands, respectively.

Table 18: Average Variance Explained (AVE) and squared latent correlations (R^2) for boredom (DUBS), work engagement (UWES), work addiction (DUWAS), and burnout (BAT) (Flanders; N = 1,500)

Scale	AVE	R^2				
		UWES	EXC	COM	DUBS	BAT
Work engagement (UWES)	.76	--	--	--	--	--
Working excessively (EXC)	.51	.07	--	--	--	--
Working compulsively (COM)	.53	.00	.59	--	--	--
Boredom at work (DUBS)	.58	.41	.12	.05	--	--
BAT core symptoms	.51	.42	.03	.15	.21	--
BAT secondary symptoms	.52	.17	.07	.23	.03	.62

Table 19: Average Variance Explained (AVE) and squared latent correlations (R2) for work engagement (UWES) and burnout (BAT) (Netherlands; N = 1,500)

Scale	AVE	R^2	
		UWES	BAT
Work engagement (UWES)	.80	--	--
BAT core symptoms	.62	.12	--
BAT secondary symptoms	.62	.07	.72

As both tables above show, the value of AVE is in most cases (much) higher than that of R^2. An exception is only observed for the two scales of the DUWAS (excessive and compulsive working) and for the core and secondary symptoms of the BAT. This means that both aspects of work addiction as well as the core and secondary symptoms of the BAT *cannot* be distinguished from each other. This is not surprising because in both cases the scales refer to a common underlying concept, namely work addiction and burnout, respectively.

Moreover, the results presented in Tables 18 and 19 confirm the discriminant validity of the BAT vis-à-vis validated scales measuring boredom at work (DUBS), work engagement (UWES), and work addiction (DUWAS). This applies to the core symptoms as well as the secondary symptoms of the BAT. In other words, the BAT measures a different concept – namely burnout – than instruments that pretend to measure work engagement , work addiction and boredom.

7. Relationships with job demands, job resources, outcomes, and personality

Based on earlier research with the Job Demands-Resources (JD-R) model, burnout – as measured by the BAT – is expected to be positively associated with stressors (job demands and work-life conflict) and negatively with job resources (Crawford, LePine & Rich, 2010; Schaufeli & Taris, 2014). The direction of the relationship between burnout on the one hand and organizational outcomes and personality on the other differs depending on the personality trait and outcome involved (see below).

Various job demands and job resources were included in the Dutch sample, whereas several personality factors were included in the Flemish sample. For the assessment of job demands and job resources the Energy Compass was used, an online questionnaire (Schaufeli, 2015, 2017b) that is based on the Job-Demands Resources model (Schaufeli & Bakker, 2004b; Bakker & Demerouti, 2016). Appendix 1 includes more detailed information about the content of the Energy Compass.

Tables 20 to 25 also include two subscales of the UBOS (i.e., exhaustion and cynicism) to allow for a comparison with similar subscales of the BAT (i.e., exhaustion and mental distance). We start with presenting correlations, whereby only those scales of the *Energy Compass* are included whose correlations with at least one burnout indicator exceeds .30. Next, the results of regression analyzes are presented that include each of the burnout scales separately as dependent variable. In doing so the unique contribution of specific job demands, job resources or personality factors is shown.

Table 20 shows the correlations between the burnout scales of the BAT and the UBOS and eight different stressors (job demands and work-life conflict). All correlations are positive and thus in the expected direction; the more job demands and work-life conflict, the more burnout symptoms. In particular, the relationships with work-life conflict[13], interpersonal conflict and role conflict are relatively strong, whereas those with bureaucracy and harassment are somewhat lower. Tellingly, the relationships with

13 Please note that the association with work-life conflict does not mean that burnout is "caused" by problems outside work. This association may be due to the fact that employees feel tired and stressed-out outside work, for instance, because of work overload or interpersonal conflict at work; that is, problems at work may have a negative impact on personal life and may therefore cause work-life conflict.

job demands that are traditionally associated with burnout – quantitative demands (work overload) and emotional job demands – are less strong. By way of comparison; a meta-analysis of several dozen studies with the MBI-GS resulted in a correlation of workload with exhaustion and cynicism of ρ = .40 and ρ = .24, respectively. (Alarcon, 2011). For role conflict, the meta-analytic correlations were. 42 and .29, respectively. In other words, compared to this meta-analysis, the correlations in Table 20 with qualitative job demands (workload) are lower (for exhaustion) or equal (for mental distance), whereas those with role conflict are higher.

Table 20: BAT and UBOS correlations (r) with job demands and work-life conflict (Netherlands, N = 1,500)[14]

Job demands	B-EX	U-EX	B-MD	U-CY	B-EI	B-CI	B-Tot	B-Sec
Qualitative job demands	.33	.29	.23	.27	.27	.29	.32	.29
Emotional job demands	.36	.33	.25	.27	.31	.28	.34	.33
Physical job demands	.32	.29	.27	.22	.25	.19	.29	.29
Bureaucracy	.47	.44	.44	.48	.45	.47	**.51**	.42
Role conflict	**.56**	**.54**	**.58**	**.60**	**.56**	**.55**	**.62**	**.51**
Interpersonal conflict	**.56**	**.53**	**.59**	**.57**	**.62**	**.59**	**.65**	**.53**
Harassment	.46	.45	.48	.45	**.54**	**.54**	**.55**	.48
Work-life conflict	**.54**	**.51**	**.50**	.48	**.55**	**.52**	**.58**	**.52**

Note: All correlations, p < .001; correlations ≥ .50 are printed in bold. B-EX = BAT exhaustion, U-EX = UBOS exhaustion, B-MD = BAT mental distance, U-CY = UBOS cynicism, B-EI = BAT emotional impairment, B-CI = BAT cognitive impairment, B-Tot = BAT total score (23 items), B-Sec = BAT secondary symptoms (10 items).

The correlations with the UBOS exhaustion scale are slightly less high than those of the BAT, while the reverse applies to the distance/cynicism scales of both instruments. However, the differences are very small and amount to at most .04. In short, broadly speaking, the BAT and the UBOS are equally strongly related to job demands and work-life conflict.

With the exception of interpersonal conflict and harassment, the exhaustion scale of the BAT is more strongly related to job demands compared to the other BAT scales. This confirms the dominant position of exhaustion as a central burnout symptom. Furthermore, there is no noticeable difference in pattern of correlations between the various scales of the BAT. However, the core symptoms (BAT total-score) are always more strongly related to job demands and work-life conflict than the secondary

[14] Correlations with negative organizational change, pace of organizational change and mental work demands are lower than .30 for all burnout scales and are therefore not displayed.

symptoms. Yet, differences are small and amount to .14 at most. Nevertheless, this again confirms the relevance of the distinction between core and secondary burnout symptoms.

Because job demands and work-life conflict are interrelated, correlations in Table 20 are inflated and do not give a proper picture of the unique contribution of each job demand individually. That is why regression analyses were carried out, which take mutual relationships into account that exist between independent variables (Table 21).

Table 21: Regression coefficients (β) of the BAT and the UBOS with job demands and work-life conflict (the Netherlands, N = 1,500)

Job demands	B-EX	U-EX	B-MD	U-CY	B-EI	B-CI	B-Tot	B-Sec
Qualitative job demands	--	--	.06*	--	--	--	--	--
Emotional job demands	.06*	--	.06*	--	--	--	--	--
Physical job demands	.12***	.11***	.09***	--	--	--	.07***	.10**
Bureaucracy	.20***	.20***	.14***	.11***	.19***	--	.20***	.21***
Role conflict	--	--	--	--	--	.07*	--	--
Interpersonal conflict	.23***	.25***	.30***	.34***	.15***	.15***	.24***	.16***
Harassment	14***	.12**	.24***	.20***	.27***	.19***	.22***	.09**
Work-life conflict	.08***	.11*	.11***	.08***	.19***	.23***	.17***	.25***
Explained variance	42%	38%	43%	41%	46%	43%	52%	40%

Note: * p <.05, ** p < .01, ** p < .001. B-EX = BAT exhaustion, U-EX = UBOS exhaustion, B-MD = BAT mental distance, U-CY = UBOS cynicism, B-EI = BAT emotional impairment, B-CI = BAT cognitive impairment, B-Tot = BAT total score (23 items), B-Sec = BAT secondary symptoms (10 items).

It appears that the exhaustion and distance scales of the BAT explain somewhat more variance in job demands and work-life conflict than the corresponding scales of the UBOS. Furthermore, the pattern of β-values hardly differs between the corresponding UBOS and BAT scales. Please note that quantitative and emotional demands are related to the distance scale of the BAT but not to cynicism of the UBOS.

Role conflict emerges as the most important job demand, whereas bureaucracy hardly plays any role, presumably due to the overlap with other job demands. It is noteworthy that work-life conflict relates to emotional but not to cognitive impairment. For the rest, the pattern of β-values differs very little between the BAT scales.

By far most variance is explained by the total score of the BAT, also compared to the secondary symptoms. In other words, the BAT-23 is best used when analyzing the relationship with job demands.

Table 22 shows an overview of eight job resources, plus engaging leadership and in-role performance. In accordance with the Job-Demands Resources model, the relationship of burnout with job resources is negative and less strong than with job demands. Team atmosphere and role clarity emerge as the most important job resources, followed by person-job fit and team effectiveness, respectively. When the team atmosphere is bad, the team is less effective, work roles are less clear and employees fit poorly with their jobs, the more burn-out complaints are experienced.

Table 22: Correlations (r) of the BAT and the UBOS with job resources, in-role performance and leadership (the Netherlands; N = 1,500)

Job resources	B-EX	U-EX	B-MD	U-CY	B-EI	B-CI	B-Tot	B-Sec
Role clarity	**-.30**	**-.32**	**-.34**	**-.36**	**-.32**	**-.32**	**-.35**	-.24
Meeting expectations	-.28	**-.32**	**-.31**	**-.34**	**-.31**	**-.31**	**-.33**	-.24
Team atmosphere	**-.35**	**-.39**	**-.41**	**-.44**	**-.37**	**-.30**	**-.40**	**-.32**
Team effectiveness	-.26	**-.30**	**-.32**	**-.35**	-.27	-.23	**-.30**	-.23
Person-job fit	-.26	**-.31**	**-.38**	**-.38**	-.21	-.22	**-.30**	-.20
Possibilities for learning	-.19	-.25	**-.33**	**-.32**	-.13	-.10	-.21	-.14
Alignment	-.23	**-.30**	**-.31**	**-.34**	-.26	-.22	-.28	-.21
Availability of tools	-.25	-.27	-.29	**-.32**	-.27	-.21	-.28	-.22
Engaging leadership	-.20	-.25	-.28	**-.32**	-.17	**-.31**	-.21	-.15

Note: All correlations, p <.001; correlations ≥ .30 are printed in bold. B-EX = BAT exhaustion, U-EX = UBOS exhaustion, B-MD = BAT mental distance, U-CY = UBOS cynicism, B-EI = BAT emotional impairment, B-CI = BAT cognitive impairment, B-Tot = BAT total score (23 items), B-Sec = BAT secondary symptoms (10 items).

The two UBOS scales are somewhat more strongly related to job resources than the corresponding BAT scales, although difference are rather small –.07 at most. Again, the core symptoms of burnout (BAT Total-score) are more strongly related than the secondary symptoms; the difference is .12 at most. In general, the correlations with emotional and cognitive impairment are somewhat less strong than with the other two BAT scales. As with job demands, the patterns of associations between each of the BAT scales and the various job resources does not really differ.

As could be expected from the correlations, the explained variance of the burnout scales is relatively small. (Table 23). This applies in particular to the secondary symptoms. The two UBOS scales explain slightly more variance than the BAT scales, although the difference of 3-4% is rather small. It is noteworthy that the explained variance of the secondary symptoms is much less than that the other burnout scales (11% vs. 20%). Again, this underlines the special nature of the secondary symptoms.

Table 23: Regression coefficients (β) the BAT and the UBOS with job resources (the Netherlands, N = 1,500)

Job resources	B-EX	U-EX	B-MD	U-CY	B-EI	B-CI	B-Tot	B-Sec
Role clarity	-.18**	-.11***	-.12***	-.13***	-.15***	-.18***	-.17***	-.08*
Meeting expectations	--	--	--	--	-.09**	-.14***	-.08**	--
Team atmosphere	-.23***	-.24***	-.24***	-.25***	-.30***	-.19***	-.28***	-.24***
Team effectiveness	--	--	--	--	-.10*	-.09*	-.09*	--
Person-job fit	-10***	-12***	-.18***	-.17***	-.07*	-.14***	-.11***	--
Possibilities for learning	--	--	-.08*	--	-.15*	-.15**	--	--
Alignment	--	-.09**	--	-.08**	-.10***	--	--	--
Availability of tools	--	--	--	--	.-06*	--	--	-.07*
Explained variance	15%	19%	23%	25%	18%	16%	20%	11%

Note: * p <.05, ** p < .01, ** p < .001. B-EX = BAT exhaustion, U-EX = UBOS exhaustion, B-MD = BAT mental distance, U-CY = UBOS cynicism, B-EI = BAT emotional impairment, B-CI = BAT cognitive impairment, B-Tot = BAT total score (23 items), B-Sec = BAT secondary symptoms (10 items).

Role clarity, team atmosphere and person-job fit emerge as the most important unique job resources that are associated with almost all burnout scales. Apart from a few minor differences, the pattern of β-values is the same for the BAT and UBOS scales. Meeting expectations, team effectiveness, and opportunities for learning and development are all related to emotional and cognitive impairment, but not to both other BAT scales.

Table 24: Correlations (r) of the BAT and the UBOS with organizational outcomes (the Netherlands, N = 1,500)[15]

Organizational outcomes	B-EX	U-EX	B-MD	U-CY	B-EI	B-CI	B-Tot	B-Sec
Organizational commitment	-.24	**-.30**	**-.36**	**-.37**	-.19	-.17	-.27	-.19
Team commitment	-.27	**-.31**	**-.37**	**-.37**	-.25	-.20	**-.30**	-.24
Job satisfaction	**-.37**	**-43**	**-.47**	**-.48**	**-.31**	-.28	**-.40**	**-.35**
Turnover intention	**.45**	**.42**	**.50**	**.51**	**.40**	**.43**	**.49**	**. 38**
Workability	**-.33**	**-.38**	-.27	**-.30**	**-.33**	-.28	**.34**	**-.33**
Work related sickness absence	**.36**	**.37**	**.33**	**.34**	**.34**	**.30**	**.37**	**.34**
In-role performance	-.27	**-.31**	**-.33**	**-.36**	**-.36**	**-.33**	**-.35**	-.28
Work performance	-.26	-.27	**-.30**	**-.30**	-.29	**-.30**	**-.31**	**-.34**

Note: All correlations, p < .001; correlations ≥ .35 are printed in bold. B-EX = BAT exhaustion, U-EX = UBOS exhaustion, B-MD = BAT mental distance, U-CY = UBOS cynicism, B-EI = BAT emotional impairment, B-CI = BAT cognitive impairment, B-Tot = BAT total score (23 items), B-Sec = BAT secondary symptoms (10 items).

[15] Correlations with duration and frequency of absenteeism, with subjective health perception and with extra-role performance are lower than .30 for all burnout scales and are therefore not included.

Table 24 provides an overview of the relationships between the burnout scales and various organizational outcomes. The pattern of correlations hardly differs between the corresponding BAT and UBOS scales. Sometimes correlations with the BAT, and sometimes those with the UBOS are higher, but the differences are very small and never exceed .07. In particular, turnover intention, work-related absenteeism, workability and job satisfaction are related to burnout. The higher the score on the burnout scales, the higher the turnover intention, the more work-related absenteeism, and the lower the levels of workability and job satisfaction.

Organizational commitment and satisfaction are more closely related to the exhaustion and distance scales of the BAT than to both other BAT scales. Apart from workability and work-related absenteeism, the correlations with the core symptoms (BAT Total-score) are higher than with the secondary symptoms.

By way of comparison, the meta-analysis by Alarcon (2011) found correlations between exhaustion on the one hand and organizational commitment, job satisfaction and turnover intention on the other of -.36, -.51, and .39, respectively. In our Dutch sample, correlations with exhaustion and organizational commitment and job satisfaction are slightly lower, and with turnover intention slightly higher, but they are all of similar size. For cynicism (mental distance), Alarcon (2011) found correlations of. -.36, -.47, and .32, respectively. In the Dutch sample, these are virtually the same for organizational commitment and job satisfaction, and somewhat higher for turnover intention. In comparison with the meta-analysis of the relationship between burn-out and performance by Taris (2006), which included 16 studies, the correlations in Table 24 for in-role performance are relatively high; this meta-analysis found correlations of UBOS exhaustion and depersonalization of -.22 and -.19, respectively.

Table 25 shows the results of the regression analyzes of the burnout scales on various outcomes. Turnover intention appears to have by far the strongest association with all burnout scales, followed at a distance by work-related absenteeism. The explained variance of the BAT and UBOS scales is about the same the predictors are also largely identical.

Table 25: Regression coefficients (β) of the BAT and the UBOS with organizational outcomes (the Netherlands; N = 1,500)

Organizational outcomes	B-EX	U-EX	B-MD	U-CY	B-EI	B-CI	B-Tot	B-Sec
Organizational commitment	--	--	-.13**	--	.15***	.10*	--	--
Team commitment	.18*	--	--	.12**	--	--	--	--
Job satisfaction	--	-20***	-.18***	-.17***	--	--	--	--
Turnover intention	.39***	.29***	.38***	.42***	.35***	.39***	.44***	.31***
Workability	-.16***	-.17***	--	.10*	-.10*	--	-.10**	-.14**
Work related sickness absence	.16***	.18***	.11**	.12**	.17***	.13**	.17***	.20***
In-role performance	--	--	--	--	-.20***	-.17***	-.10*	--
Work performance	--	--	-.09*	--	-.11*	-.12**	-.09*	-.09*
Explained variance	31%	32%	40%	43%	31%	29%	37%	24%

Note: * p <.05, ** p < .01, ** p < .001. B-EX = BAT exhaustion, U-EX = UBOS exhaustion, B-MD = BAT mental distance, U-CY = UBOS cynicism, B-EI = BAT emotional impairment, B-CI = BAT cognitive impairment, B-Tot = BAT total score (23 items), B-Sec = BAT secondary symptoms (10 items).

Notably, in-role performance and overall work performance (rated between 1-10) are negatively related to both types of BAT impairment, but not to the other two BAT scales. Here too, the secondary symptoms explain less variance than the core symptoms, although the pattern of β-values does not differ much.

Table 26 shows the correlations in the Flemish sample of the burnout scales with personality, including the Big Five personality traits. By far the highest correlations are found for neuroticism and optimism: the less emotionally stable and the less optimistic, the higher the burnout scores. Of the other Big Five traits only conscientiousness is somewhat related to burnout; the less conscientious, the more burnout complaints.

Table 26: BAT correlations with personality (Flanders; N = 1,500)[16]

Personality trait	B-EX	U-EX	B-MD	U-CY	B-EI	B-CI	B-Tot	B-Sec
Neuroticism	**.49**	**.48**	**.38**	**.38**	**.45**	**.40**	**.51**	**.59**
Conscientiousness	-.27	-.15	-.24	-.19	-.14	**-.37**	-.28	-.09
Self-efficacy	-.24	-.22	-.29	-.27	-.22	**-.30**	**-.30**	-.19
Optimism	**-.52**	**-.53**	**-.55**	**-.58**	**-.43**	**-.39**	**-.57**	**-.48**
Resilience	**-.32**	-.29	**-.32**	**-.31**	**-.31**	**-.34**	**-.38**	**-.30**

Note: All correlations, p <.001; correlations ≥ .30 are in printed bold. B-EX = BAT exhaustion, U-EX = UBOS exhaustion, B-MD = BAT mental distance, U-CY = UBOS cynicism, B-EI = BAT emotional impairment, B-CI = BAT cognitive impairment, B-Tot = BAT total score (23 items), B-Sec = BAT secondary symptoms (10 items).

[16] Correlations with extraversion, openness, friendliness, flexibility and self-control are lower than .30 for all burnout scales and are therefore not included in the table.

The correlations in Table 26 largely correspond to those from a meta-analysis by Alarcon, Eschleman and Bowling (2009). For exhaustion they found correlations with neuroticism (*r* = .50), conscientiousness (*r* = -.26), self-efficacy (*r* = -.33), and optimism (*r* = .26). The size of the first two correlations correspond to those from Table 26, but compared to the meta-analysis the correlations with self-efficacy and optimism in the Flemish sample are lower and higher, respectively. For depersonalization, meta-analytic correlations were found of . -.40, -.26, -.24 and -.27, respectively. Apart from optimism (-.27), these values are quite similar to those displayed in Table 26 for mental distance/cynicism.

The correlations with the two corresponding scales of the BAT and the UBOS hardly differ, if at all. Only the correlation between conscientiousness and exhaustion differs, which is somewhat (.12) stronger for the BAT than for the UBOS. The pattern of correlations does not differ much between the BAT scales, though.

Table 27: Regression coefficients (β) of the BAT and the UBOS with personality factors (Flanders N = 1,500)

Personality trait	B-EX	U-EX	B-MD	U-CY	B-EI	B-CI	B-Tot	B-Sec
Neuroticism	.35***	.27***	.11***	.09**	.31***	. 21***	.28***	.48***
Conscientiousness	-.08***	--	-.11***	-.05*	--	-.27***	-.13***	-.06*
Self-efficacy	-.34***	-.38***	-.47***	-.53***	-.27***	-.19***	-.39***	-.24***
Optimism	--	--	--	--	--	-.07*	--	--
Explained variance	34%	33%	32%	35%	25%	27%	40%	39%

Note: * p <.05, ** p < .01, ** p < .001. B-EX = BAT exhaustion, U-EX = UBOS exhaustion, B-MD = BAT mental distance, U-CY = UBOS cynicism, B-EI = BAT emotional impairment, B-CI = BAT cognitive impairment, B-Tot = BAT total score (23 items), B-Sec = BAT secondary symptoms (10 items).

As Table 27 shows, self-efficacy and resilience hardly play a role as predictors of burnout. In contract, neuroticism and optimism are the main predictors, and conscientiousness also matters to some extent. The pattern of β-values is almost identical for the corresponding BAT and UBOS scales. Also the predictors of the BAT scales are largely the same, except that conscientiousness plays the most prominent role in cognitive impairment. The fact that conscientious employees suffer relatively little from cognitive impairment is plausible and strengthens the validity of this BAT scale.

8. The general version of the BAT

In addition to the work-related version, the general version of the BAT was also administered online among 512 Flemish employees (79.7% women, 20.3% men; mean age 39.2 years, SD = 10.0). Respondents were approached via a Facebook group of practitioners who, in turn, forwarded the questionnaire to their clients who suffered from burnout. Of this group, 18% (N = 94) was on sick leave. Over three-quarters (67%) indicated that their sickness absence is job related. Table 28 provides an overview of some characteristics of both groups.

Table 28: Characteristics of the Flemish group of working employees (N = 416) and those currently sick listed (N = 96)

	Working	**Sick listed**
Male	21.2%	17%
Female	78.8%	83%
Average age (SD)	38.5 (10.2)	42.4 (9.2)
Frequency of sickness absence - times/yrs. (SD)	1.5 (1.6)	17.3 (72.7)
Duration of sickness absence – days/yrs. (SD)	13.9 (32.3)	139.4 (129.6)
Treated for burnout	19%	62.5%

As can be seen from Table 28, the sick listed group has been on sick leave longer and more often than those who are working. Moreover, they have been treated more frequently for burnout complaints during the last 5 years. All in all, this means that the sick listed group is considerably less healthy than the representative Flemish sample, whose average sickness absence is 10 days per year (SD = 34.9) and frequency of absence is 1.04 per year (SD = 1.5), and of whom only 6.5% has been treated for burnout.

Because the work-related version of the BAT cannot be filled out by those who are sick listed and currently not working, it was decided to split the total group into those who work (N = 416) and those on sick leave (N = 96). As expected, those on sick leave score significantly higher on all general BAT scales than those who are currently at work: exhaustion ($t_{(510)}$ = 9.50, p <.001), mental distance ($t_{(510)}$ = 5.19, p < .001), emotional impairment ($t_{(510)}$ = 7.56, p <.001), cognitive impairment ($t_{(510)}$ = 6.16, p < .001), BAT core symptoms ($t_{(510)}$ = 9.26, p < .001), and BAT secondary symptoms ($t_{(510)}$ = 16.67, p < .001).

This means that those who are on sick leave are more likely to suffer from burnout compared to those who are not on sick leave.

Table 29: Means (M), standard deviations (SD) and internal consistencies (coefficient α) of the general version of the BAT among working and sick listed employees in Flanders.

Scale	# items	Working (N=4136)			Sick listed (N=96)		
		M	SD	α	M	SD	α
Exhaustion	8	3.44	.89	.96	4.25	.71	.95
*Mental distance**	4	2.81	1.02	.91	3.40	.93	.83
Emotional impairment	5	3.09	.82	.95	3.85	.95	.94
Cognitive impairment	5	3.20	.93	.95	3.99	.85	.96
Total core symptoms	22	3.22	.78	.97	3.95	.67	.95
Secondary symptoms	10	3.34	.74	.90	3.95	.65	.90

Note: * Identical to the work-related version of the BAT, but without the item "At work, I do not think much about what I am doing and I function on autopilot".

It appears that the correlations between the work-related BAT and the general version very high: exhaustion (r = .85) mental distance (r = .91) emotional impairment (r = .95), cognitive impairment (r = .95), and BAT core symptoms (r = .97). This means that the corresponding scales overlap between 72-94%.

In a group of 964 Flemish respondents who were approached online[17], the general version of the BAT was administered three times with an interval of 6 months (for more details about this sample, see under 5.2, p. 49).

Table 30: Test-retest reliability of the general BAT (stability-coefficient r_t)

Scale	T1 -T2 (N = 597)	T2-T3 (N = 368)	T1-T3 (N=447)
Exhaustion	.76	.78	.71
Mental distance	.68	.63	.60
Emotional impairment	.69	.68	.66
Cognitive impairment	.72	.73	.66
Total core symptoms	.78	.79	.72

The stability coefficients of the general BAT subscales range from .63 to .78 over a 6-month period, and from .60 to .71 over a 1-year period (Table 30). Furthermore, it

[17] Using the website http://burn-out.vlaanderen

appears that the test-retest reliability of the total BAT (core symptoms) is higher than that of the individual subscales. The stability coefficients therefore hardly differ from those of the work-related version (see Table 12, p. 50).

8.1. Factor structure

Four factors emerged from an exploratory factor analysis that included all work-related *and* general items of the BAT. Each of these factors included all items of corresponding work-related and general BAT scales: (1) exhaustion (23% explained variance, factor loadings vary from .54 to .79), (2) emotional impairment (18% explained variance, factor loadings range from .60 to .79), (3) cognitive impairment (18% explained variance, factor loadings range from .65 to .77), and (4) mental distance (13% explained variance, factor loadings range from .65 to .78). Furthermore, the first unrotated factor explains 56% of the variance and the factor loadings vary from. 61 to .84. This suggests that the BAT measures one underlying concept, from which four closely related aspects can be distinguished.

The very high correlations between the two versions of the BAT and the fact that the items of both versions cluster into similar rather than different work-related and general factors, indicate that the general version is, in fact, identical with the work-related version of the BAT.

Subsequently, using confirmatory factor analyses, it was investigated to what extent three different models fit the data. The 1-factor model assumes that all items of the general version load on a one, single factor (Model 1). The 4-factor model assumes four correlated factors: exhaustion, mental distance, and emotional and cognitive impairment (Model 2). Finally, the second-order model assumes that these four factors all load on one underlying latent factor (Model 3). This model corresponds to the idea that burnout is a syndrome that consists of four interrelated symptoms that refer to one underlying condition.

The three models were tested simultaneously in the group of working and sick listed employees, using the multiple group method. Table 31 shows the fit indices of the three models.

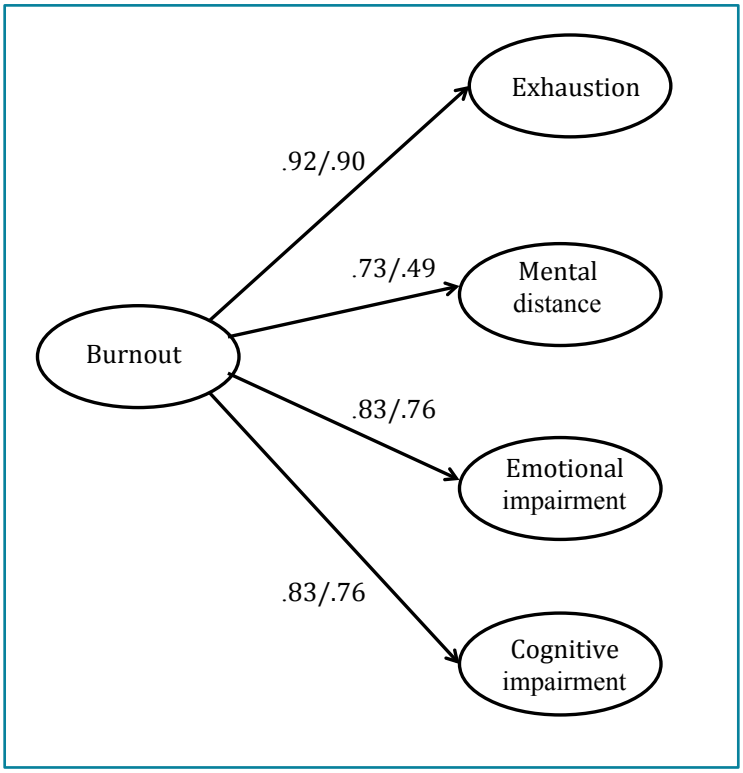
Table 31: Confirmatory factor analysis of the general version of the BAT (working N = 416; sick listed N = 96)

Model		χ^2	df	CFI	TLI	RMSEA
1	*1-factor model*	4102.16	418	.69	.66	.13
2	*4-factor model*	1033.70	406	.95	.94	.05
3	*Second-order model*	1038.70	410	.95	.94	.05

Note: χ^2 = chi-square, df = degrees of freedom; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation.

The 1-factor model (Model 1) does not fit, but both the 4-factor model (Model 2) and the second-order model (Model 3) both fit well to the data. The fit of Model 2 and Model 3 does not differ significantly: $\Delta\chi^2 = 5$, df = 4, n.s. However, the second-order model is preferred because it corresponds to the basic idea underlying the BAT, namely that burnout is a syndrome. Figure 6 shows the second-order factor loadings.

Figure 6: The structure of the general version of the BAT



Note: working/sick listed.

The second-order loadings are somewhat lower in the sick listed group. This applies in particular to mental distance. Apparently, compared to the other aspects of burnout mental distance does not play a major role amongst those who are currently on sick leave. This is unsurprising, because it is very likely that work is at the top of their minds.

8.2. Validity

To investigate whether the sick listed group scores – as expected – unfavorably on various indicators of health and well-being, scores on the general version of the BAT have been related to:

- Job satisfaction: "To what extent are you, overall, satisfied with your job?" Scoring: 1 = very dissatisfied, 2 = dissatisfied, 3 = not dissatisfied/not satisfied, 4 = satisfied, 5 = very satisfied.
- Subjective health: "What do you generally think of your health?" Scoring: 1 = poor, 2 = moderate, 3 = good, 4 = very good, 5 = excellent.
- Happiness: "On a scale of 0 -10, how happy do you feel?"
- Feeling burned out: "Do you feel burned-out or drained by your work?" Scoring: 1 = never, 2 = rarely, 3 = occasionally, 4 = often, 5 = always.
- Depression: Beck Depression Inventory II (Beck, Steer & Gregory, 1996; 21 items, α = .88)

Table 32: Correlations (r) of the general version of the BAT with various indicators of health and well-being (N = 96)

Scale	Job satisfaction	Subjective health	Happiness	Burned-out	Depression (BDI)
Exhaustion	-.37**	-.46**	-.50**	.66**	.50**
Mental distance	-.62**	-.01	-.30**	.39**	.31**
Emotional impairment	-.13	-.15	-.46**	.44**	.48**
Cognitive impairment	-.27**	-.25**	-48**	.44**	.37**
Total core symptoms	-.43**	-.30**	-.56**	.62**	.53**
Secondary symptoms	-.31**	-.32**	-.46**	.54**	.51**

Note: ** $p < .01$.

All correlations shown in Table 32 are in the expected direction, although three are not significant. Exhaustion correlates moderately to strongly with all indicators. Mental distance is particularly strongly related with job satisfaction, but not with subjective health. Emotional impairment is most strongly associated with depression (positive) and happiness (negative), whereas cognitive impairment is particularly related with happiness (negative) and with feeling burned-out (positive). The overall score on the BAT as well as the secondary symptoms are moderately related with *all* indicators. That subjective health is not so strongly related with the overall BAT score might be

explained by the fact this refers to physical rather than mental health. The strongest relationships of the BAT are found with indicators of *mental* health; that is, depression and the feeling burned out.

All in all, the correlations from Table 32 confirm the validity of the general version of the BAT. That means that those who score high on the BAT are not satisfied with their work and do not feel very healthy and happy, but burned-out and depressed instead.

9. Conclusions

A number of psychometric analyzes have been carried out to validate the BAT, mainly using independent, representative samples from the Flemish (N = 1,500) and Dutch (N = 1,500) working population. This led to the following conclusions.

The scores on two core dimensions of BAT – exhaustion and cognitive impairment – are approximately normally distributed. This is not the case with mental distance and emotional impairment, where relatively many low scores are found. The picture of the secondary symptoms is mixed; some are normally distributed, whereas others are not. It is not surprising that high item scores occur relatively rarely, because in principle a healthy group of employees was sampled. Moreover, the observed minor deviations from normality do not obstruct the use of statistical analyzes.

Although four core aspects of burnout can be distinguished, they are so strongly interconnected that one can speak of a single burnout construct. This is illustrated by the fact that a second-order factor model, in which burnout is considered as a syndrome consisting of four core symptoms, fits well with the data from both samples. This means that in addition to scores for each core dimension seperately, the total BAT score can also be used.

Both types of secondary burnout symptoms (i.e. psychological distress and psychosomatic complaints) are so closely related that it makes little sense to distinguish between them. Therefore, both are combined into one single scale that taps psychosomatic and psychological distress. One unsound item has been removed ("My weight fluctuates without being on a diet"), so the total scale consists of 10 items. Moreover, psychosomatic and psychological distress can be distinguished from depressive symptoms.

Core and secondary burnout symptoms are strongly interrelated and share 50-64% of their variance, depending on the sample. Moreover, a second-order factor analysis that comprises all BAT scales (including depression) yields one single factor. As expected, exhaustion loads highest on this factor. Again, burnout appears to be one concept, even when the secondary symptoms are considered.

Based on the scoring on the BAT scales, 6 groups or clusters can be distinguished that differ in level of burnout, varying from very low to very high. The validity of this distinction was demonstrated because employees in the higher scoring groups have more health complaints, experience higher job demands and have fewer job resources at their disposal. It is important to note that there are no groups that score high on one scale and low on another or vice versa. In other words, if someone is classified "high", "low", etc. on one subscale then that applies likewise to all other BAT subscales. This is an additional argument for using the total score on the BAT as an indicator for burnout.

The internal consistency (Cronbach's α) of the BAT scales is excellent. The test-retest reliability is also good; on the one hand the BAT scores are sufficiently stable over a period of 6 months, yet on the other hand, they also show ample variation. Although the inter-rater reliability of the BAT is low to moderate, it is better than that of the MBI with the exception of cognitive impairment. Apparently self-reported burnout differs somewhat from peer ratings by one's partner.

On the basis of a Multi-Trait, Multi-Method analysis, the convergent validity of the BAT with respect to the UBOS and the OLBI was investigated. It appears that the exhaustion and mental distance scales of the BAT largely overlap with the corresponding scales of these two other burnout instruments. At the same time, however, burnout, as measured by the four core dimensions of the BAT, can be distinguished from the way burnout is measured by the UBOS (exhaustion, cynicism, and reduced efficacy) and the OLBI (exhaustion and disengagement). In summary: the BAT partially overlaps with corresponding burnout dimensions that are also measured by other questionnaires (convergent validity), but it also adds something unique (discriminant validity).

Burnout, as assessed by the BAT can be distinguished from boredom at work (as measured by the DUBS), work engagement (as measured by the UWES), and work

addiction (as measured by the DUWAS). In other words, burnout cannot be reduced to other types of work-related well-being. Moreover, it also appears that the core symptoms of the BAT cannot be distinguished from the secondary symptoms, which is compatible with the idea of a singe burnout construct.

The core dimensions of the BAT as well as its total-score are related to a large number of job demands, job resources, personality factors and organizational outcomes. More specifically it was observed that:

- The BAT is positively associated with job demands, in particular with work-life conflict, interpersonal conflict, role conflict, bureaucracy, and harassment.
- The BAT is negatively related to a number of job resources, such as team spirit , role clarity, and person-job fit. As expected on the basis of the Job Demands-Resources model, these associations are less strong than with job demands.
- The BAT is negatively related to job satisfaction, turnover intention, workability and work performance, and positively to sickness absenteeism.
- In particular, the BAT is positively associated with neuroticism and negatively with optimism. Negative associations with resilience, conscientiousness and self-efficacy are somewhat lower.
- With one exception (neuroticism) correlations of job demands, job resources, personality factors and organization outcomes are stronger with the core symptoms of the BAT than with the with the secondary symptoms. This is in line with the rationale that lies behind this distinction.
- Overall, the pattern of associations of job demands, job resources, organizational outcomes and personality factors with the BAT hardly differs from that of the UBOS, if at all. Sometimes correlations with the BAT are slightly stronger, but overall differences are very small and do not exceed .07.
- Generally, the scales of the BAT are all in a similar way related to job demands, job resources, personality factors and organization outcomes, although exhaustion sometimes shows a somewhat stronger association.
- Usually, associations with the total-score of the BAT are stronger than with the individual subscales. This makes sense because the total score reflects *all* core aspects of burnout.

The general conclusion is that the BAT is a reliable and valid questionnaire to measure burnout.

For psychometric research with the general version of the BAT, an online sample was used consisting of 416 employees, who also completed the work-related version, as well as 96 employees who are currently on sick leave. Compared to the former group the latter group has significantly higher scores on all BAT scales. Other important findings are:

- The general BAT scales have a very good internal consistency.
- The correlations between the work-related and general scales of the BAT are very high. In addition, work-related and general BAT-items constitute joint factors for each scale. Both findings illustrate that the items of both versions overlap so strongly that it can be assumed that they measure the same concept.
- Confirmatory factor analyzes show that a second-order model, in which burnout is seen as a syndrome of four strongly related factors that refer to one underlying condition, fits the data well. The fact that all items load high on the first unrotated factor that accounts by far for the most variance, points in the same direction. In practical terms, this means that in addition to the four scale scores, the overall score of the general BAT can also be used.
- The validity of the general version of the BAT is supported by the moderate to strong associations with a number of indicators for health and well-being, such as job satisfaction, subjective health, happiness, the feeling of being burned-out, and depression, as measured by the BDI.

The conclusion is that the general version of BAT is a reliable and valid questionnaire to measure burnout outside the work context.

PART III: THE SHORT VERSION OF THE BAT

In practice in particular, there is a great need for short, reliable and valid questionnaires so as not to burden respondents unnecessarily. We opted for a two-stage approach, in which a relatively long questionnaire consisting of 23 items was developed first to ensure that no aspects of burnout would be overlooked. This comprehensive questionnaire was then shortened and its psychometric qualities (i.e., reliability and validity) were investigated. The results are described in this part of the manual, whereby also a comparison is made with the original, longer version. Ideally, after shortening the reliability and validity of BAT should not deteriorate.

Part III discusses the selection and frequency distribution of the items of the shortened BAT, its factorial validity, its relationship with the BAT-23, its reliability, convergent and discriminant validity, and its relationships with job demands, job resources, organizational outcomes and personality (content validity). In addition to the work-related version of the BAT, a general version of the shortened BAT is also presented.

1. Item selection

The aim of shortening the BAT-23 is to arrive at a shorter questionnaire that takes into consideration the complexity of the burnout syndrome. This means that the short version must also consist of four core symptoms, each of which is measured using a separate subscale. In such a case, for simplicity and transparency, it is recommended that the subscales should consist of the same number of items (Nielsen & Kreiner, 2011). The minimum number of items per subscale is set at three, so that the short version of the BAT consists of 12 items (BAT-12).

Rasch analysis was used to select the items for the short version, whereby various indicators were used as criteria for item selection[18]. Rasch analysis is a psychometric technique based on item response theory, which is often used for the development

[18] We wish to acknowledge Emina Hadzibajramovic, PhD (Gothenburg University) for her help in carrying out the Rasch analyzes.

and psychometric evaluation of questionnaires[19]. The purpose of Rasch analysis is to determine whether the observed data meet the requirements of the Rasch model. If that is the case, the questionnaire has solid psychometric properties. The advantage of the Rasch model over classic test theory approaches, such as factor analysis, is that there is no need for normally distributed items. Therefore, Rasch analysis is preferable in the case of ordinal data, which are used in questionnaires with ordered categorical answers, such as the BAT, that is scored on an ordinal frequency scale ranging from 1 (never) to 5 (always).

Items that perform poorly on one or more item fit-indicators of the Rasch analysis are considered potential candidates for elimination. The following four item fit-indicators were used: (1) the discriminative power of the item (based on the item fit residues and the $\chi 2$ value); (2) the appropriateness of the response categories (based on the threshold order); (3) the independence of the response from other items (based on the residual correlations) and; (4) the absence of differential item functioning for age, gender and country (DIF). Absence of DIF means that items, given the same degree of burnout, function in the same way for all comparable groups (women and men, younger and older age groups, Flanders and the Netherlands).

In addition to the Rasch analysis, content analysis of the items was also used for item selection (Nielsen & Kreiner, 2013). Based on the results of content analysis, the items can be divided into five groups: (1) unproblematic items; (2) items formulated in a problematic way; (3) items whose wording overlaps with that of one or more other items; (4) items that measure the same characteristic as one or more other items; and (5) items that do not properly reflect the construct.

Based on a combination of the results of Rasch analysis and content analysis, certain items have been eliminated one by one – per subscale – according to an iterative procedure. That is, a Rasch analysis was performed again after removal of a certain item. For the purpose of cross-validation, analyses were conducted independently using two randomly drawn samples of 800 respondents each from the validation samples, which included both Flemish and Dutch employees (see Table 4; pp. 37-38).

[19] For a brief introduction please see Hagquist, Bruce and Gustavsson (2009) and for a more comprehensive overview, see the textbook of Christensen, Kreiner and Masbah (2013).

As a result, the following 12 items emerged which together constitute the shortened version of the BAT.

Table 33: The short version of the BAT

Exhaustion

1. At work, I feel mentally exhausted
2. After a day at work, I find it hard to recover my energy
3. At work, I feel physically exhausted

Mental distance

4. I struggle to find any enthusiasm for my work
5. I feel a strong aversion towards my job
6. I'm cynical about what my work means to others

Emotional impairment

7. At work, I feel unable to control my emotions
8. I do not recognize myself in the way I react emotionally at work
9. At work I may overreact unintentionally

Cognitive impairment

10. At work, I have trouble staying focused
11. When I'm working, I have trouble concentrating
12. I make mistakes in my work because I have my mind on other things

2. Distribution of scores

Table 34 shows the distribution characteristics of the BAT-12 items. As noted above (p. 39), due to the size of the samples, it makes no sense to perform a formal test for normality of the items. Instead, a visual inspection of the frequency distributions of the items As can be seen from Table 34 the exhaustion and cognitive impairment items are fairly normally distributed, with a score of 2 being most common and scores 4 and 5 the least common. On the other hand, the distribution of mental distance and emotional impairment scores differs somewhat from a normal distribution; the lowest score of 1 occurs most often and the frequency of scores decreases successively as

values increase from 2 to 5. Overall, the pattern of the distribution of the items is comparable for the Netherlands and Flanders.

Table 34: Range of the means (M), standard deviations (SD), skewness and kurtosis of the BAT-12-items for Flanders (N = 1,500) and the Netherlands (N = 1,500)

BAT-12	M	SD	Skweness	Kurtosis
Flanders	1.66 – 2.31	.74 – 1.02	.51 – 1.16	-.46 – .88
The Netherlands	1.99 – 2.37	.95 –1.10	.36 – .82	-.57 – .25

It can be concluded that the scores on the BAT-12's exhaustion and cognitive impairment items are approximately normally distributed. However, this is not the case for the other two core symptoms; mental distance and emotional impairment. Respondents are generally less affected by these burnout symptoms. It should be noted that the fact that the score distribution of some BAT items deviates from normality, does not obstruct the use of the data analyzis techniques performed below, as these are fairly robust for violations of normality.

3. Factorial validity

Both exploratory and confirmatory factor analysis of the BAT-12 items have been performed. In the Dutch sample (N = 1,500), an explorative factor analysis yields a single factor that explains 65% of the variance and on which all items load at least .75. In the Flemish sample (N = 1,500) three factors emerged that explain 28%, 23% and 22% of the variance, respectively (together 73%). The first factor includes all exhaustion and mental distance items (factor loadings range from .52 to .84), the second factor includes the three emotional impairment items (loadings range from .80 to .85), and the third factor includes all cognitive impairment items (loadings range from .78 to .84). In addition, a double loading of the exhaustion item "*At work I feel mentally exhausted*" is observed on the second factor (cognitive impairment).

In summary: the result of the explorative factor analysis differs between the Netherlands and Flanders. Only one factor is found in the Netherlands, while three factors emerge in Flanders, which correspond to the four different aspects of burnout, with the proviso that the items related to exhaustion and mental distance cluster into one factor.

Furthermore, three models were tested using confirmatory factor analysis: a 1-factor model in which all 12 items are expected to load on one underlying factor (Model 1); (2) a 4-factor correlated model, in which the four core aspects (exhaustion, mental distance, and cognitive and emotional impairment) are distinguished and correlated with each other (Model 2); (3) a second-order model, assuming that the four core aspects refer to one underlying burnout factor (Model 3). Hence, Model 3 is consistent with the view that burnout is a syndrome consisting of four symptom dimensions.

The critical values for the fit indices in the Table 35 are (see also p. 41): *Comparative Fit Index* (CFI $\geq$ .90), *Tucker-Lewis Index* (TLI $\geq$ .90) and the *Root Mean Square Error of Approximation* (RMSEA $\leq$ .08). In addition to simultaneous testing (using the multiple group method), the three models were also tested separately in the Flemish and Dutch sample because the explorative factor analysis yielded different results.

As shown in Table 35, the 1-factor model does not fit well with the data in both Flanders and the Netherlands. The 4-factor and the second-order model, on the other hand, both fit about equally well. The same applies if the three models are tested simultaneously on both samples.

Table 35: Confirmatory factor analysis of the BAT-12 (Flanders, N = 1,500; the Netherlands, N = 1,500; both samples, N = 3,000)

Model		χ^2	df	CFI	TLI	RMSEA
	Flanders					
1	*1-factor model*	3311.75	54	.72	.66	.20
2	*4-factor model*	221.21	48	.99	.98	.05
3	*Second-order model*	222.23	50	.99	.98	.05
	The Netherlands					
1	*1-factor model*	2177.43	154	.86	.83	.16
2	*4-factor model*	135.62	84	.99	.99	.04
3	*Second-order model*	201.24	50	.99	.99	.05
	Both samples					
1	*1-factor model*	5489.19	108	.80	.76	.13
2	*4-factor model*	356.83	96	.99	.98	.03
3	*Second-order model*	423.46	100	.99	.98	.03

Note: χ^2 = chi-square df = degrees of freedom; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation.

So despite the fact that only one factor emerged from the explorative factor analysis in the Dutch sample, it appears that here too four aspects of burnout can be distinguished. However, these four aspects are very closely related, as shown in Table

36. In the Dutch sample, the factors are on correlated on average .80 compared to .65 in the Flemish sample.

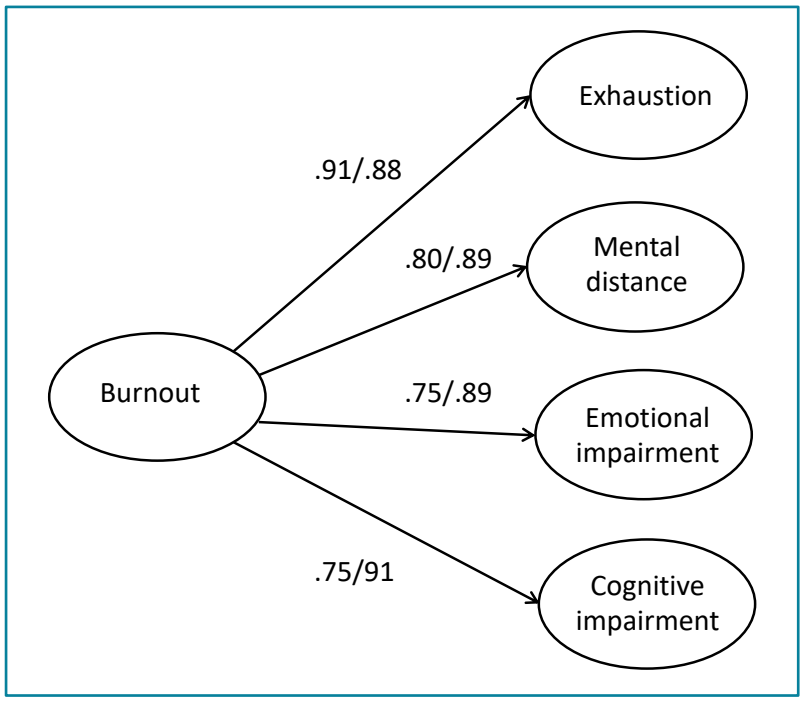
Table 36: Correlations (r) between the latent BAT-12 factors and the BAT-12 subscales in Flanders (N =1,500) and the Netherlands(N = 1,500)

BAT-12		1	2	3	4
1	*Exhaustion*	--	.84 (.74)	.76 (.68)	.79 (.69)
2	*Mental distance*	.72 (.63)	--	.78 (.71)	.79 (.71)
3	*Emotional impairment*	.69 (.59)	.60 (.53)	--	.84 (.77)
4	*Cognitive impairment*	.67 (.59)	.61 (.53)	.58 (.51)	--

Note: Values for the Netherlands are displayed above the diagonal and for Flanders below the diagonal. The observed, manifest correlations between the subscales are displayed between brackets. All correlations are significant (p <.001).

The strong interrelation between the different aspects of burnout is also evident in the second-order model (Model 3). Figure 7 shows to what extent the four aspects refer to the same underlying factor, which can be interpreted as burnout.

Figure 7: The structure of the BAT-12



Note: Flanders/the Netherlands

As expected based on the correlations from Table 36, the factor loadings on the second-order factor in Flanders are lower than in the Netherlands. Please note that the correlations between the latent factors are *by definition* higher than the observed, manifest correlations between the subscales because the former are free of measurement errors. The pattern of correlations between the latent factors on the one

hand and the observed correlations on the other hand is virtually the same, with only small differences ranging between .07 and .10.

In practical terms, the BAT-12 can be interpreted both as a unidimensional questionnaire that measures burnout on the basis of a single total score, *and* as a four-dimensional questionnaire that measures different aspects of burnout, which are strongly interrelated. In the remainder of this manual, therefore, the total score is used as well as the scores on the four subscales of the BAT-12.

4. Relationship between the full and short versions

How strongly relates the shortened version (BAT-12) to the full version (BAT-23)? Tables 37 and 38 give an overview of the correlations in the Dutch and Flemish sample between the subscales of both versions of the BAT, as well as of the correlations with the total scores and secondary symptoms.

Table 36: Correlations between the BAT-12 and the BAT-23 (the Netherlands; N = 1,500)

		BAT-12				
BAT-23		**1**	**2**	**3**	**4**	**5**
1	***Exhaustion***	**.96**	.77	.75	.74	.89
2	***Mental distance***	.74	**.97**	.71	.72	.89
3	***Emotional impairment***	.70	.74	**.98**	.78	.90
4	***Cognitive impairment***	.72	.73	.79	**.98**	.90
5	***Total BAT-12***	.87	.90	.89	.89	1.00
6	***Total BAT-23***	.88	.89	.87	.88	**.99**
7	***Secondary symptoms***	.72	.66	.71	.71	.79

Note: Correlations between the corresponding scales of the BAT-23 and BAT-12 are shown in bold.

First of all, it appears that the total scores of both versions of the BAT correlate almost perfectly with each other (r = .99), while more or less the same applies to its corresponding subscales (.96 < r <.98). This means that the (sub) scales are in fact identical. Furthermore, the scores on the BAT-12 subscales correlate very strongly with the total score; this is true for the BAT-12 as well as the BAT-23. This in yet another indication that the subscales of the short BAT assess the same construct as the corresponding subscales of full version. Finally, the correlations of the BAT-12 subscales with the secondary burnout symptoms are in the same order as those of the BAT-23 from Table 10 (p. 45). The differences are extremely small and amount to .03 at most.

Also in Flanders it appears (Table 38) that the total scores of both versions of the BAT correlate almost perfectly with each other (r = .99). Likewise this applies for the corresponding subscales (.96 < r < .98). Moreover, the subscales of the BAT-12 also correlate high with the total scores of both versions of the BAT, albeit slightly less than in the Netherlands. Finally, the correlations of the BAT-12 with secondary symptoms and depressed mood – as measured by the depression scale of the 4DKL – hardly differ from those with the BAT-23. In fact, the maximum difference in correlations is .04 for secondary burn-out complaints and .03 for depressed mood.

Table 38: Correlations between the BAT-12 and the BAT-23 (Flanders; N = 1,5

		BAT-12				
BAT-23		1	2	3	4	5
1	*Exhaustion*	**.96**	.69	.61	.65	.89
2	*Mental distance*	.60	**.97**	.51	.54	.81
3	*Emotional impairment*	.63	.58	**.97**	.54	.82
4	*Cognitive impairment*	.62	.54	.53	**.98**	.80
5	*Total BAT-12*	.86	.84	.79	.79	1.00
6	*Total BAT-23*	.86	.83	.76	.79	**.99**
7	*Secondary symptoms*	.70	.52	.56	.53	.70
8	*Depressed mood*	.53	.53	.49	.42	.60

Note: Correlations between the corresponding scales of the BAT-23 and BAT-12 are shown in bold.

Conclusion: The correlations between the corresponding subscales of the BAT-12 and BAT-23 and those of the total scores are almost identical in both samples, indicating that they measure the same construct. This is confirmed by the same pattern of correlations that both versions of the BAT show with regard to secondary complaints and depressed mood.

5. Reliability

Two aspects of reliability were considered: internal consistency and test-retest reliability over a period of six and twelve months, respectively.

5.1. Internal consistency

Table 39 shows the values of Cronbach's α, which indicate to what extent the scale is internally consistent. In general, a values $\geq$ .70 are considered sufficient and values $\geq$ .80 are considered good (Nunally & Bernstein, 1994).

Table 39: Internal consistency of the BAT-12 (coefficient α)

BAT-12	# items	Flanders (N = 1,500)	Netherlands (N = 1,500)
Exhaustion	3	.85	.87
Mental distance	3	.88	.90
Emotional impairment	3	.87	.92
Cognitive impairment	3	.87	.90
Total BAT-12	12	.92	.95

All subscales of the BAT-12 as well as the total score have a very good internal consistency in both samples. Values of α are by definition lower than for the BAT-23 (see Table 11, p. 49) because α depends on the number of items of the scale; the fewer items, the lower the value of α. For that reason, the difference between the exhaustion scales is the largest (.07), as it consists of 8 items, while the differences between the other three subscales, which each consist of 5 items, are less (ranging from .02 to .05).

5.2 Test-retest reliability

The BAT was administered three times with a time interval of 6 months in a group of 964 Flemish respondents who were approached online (for more details see p. 49-50). Table 40 summarizes the test-retest reliability using auto-correlations or stability coefficients.

Table 40: Test-retest reliability of the BAT-12 (stability coefficient r_t)

BAT-12	T1 -T2 (N = 597)	T2-T3 (N = 368)	T1-T3 (N=447)
Exhaustion	.65	.70	.63
Mental distance	.68	.65	.59
Emotional impairment	.68	.64	.57
Cognitive impairment	.60	.55	.50
Total BAT-12	.74	.72	.67

Note: For all r_t's, $p < .0001$

The stability coefficients of the BAT-subscales range from .55 to .70 over a 6 month period, and from .50 to .63 over a 1 year period. Furthermore, it should be noted that the test-retest reliability of the total BAT-12 is higher than that of its subscales. The values from Table 40 correspond to those of the BAT-23 (see Table 12, p. 50); the differences are at most .06 for the subscales and .01 for the total score.

It can therefore be concluded that the test-retest reliability of the shortened BAT does not differ from that of the longer version. The stability coefficients of the subscales of the BAT-12 show that approximately 40% of the variance remains stable over a 6 month period and approximately 35% over a 1 year period. For the total BAT-12 that is about 50% and about 45%, respectively.

6. Convergent and discriminant validity

It has been investigated to what extent the BAT-12 compares to two other burnout questionnaires, whereby it is of course expected that the BAT-12 is positively and substantially related to these alternative burnout instruments (convergent validity). In addition, it was investigated to what extent the BAT-12 is related to questionnaires that measure other concepts such as work engagement, boredom and work addiction. In this case it is expected that there will be relatively little overlap because of the different nature of these concepts (discriminant validity).

6.1. Other burnout questionnaires

To establish the convergent validity of the BAT-12, correlations were calculated with the two most commonly used burnout questionnaires, the Utrecht Burnout Scale (UBOS; Schaufeli & Van Dierendonck, 2000), being the Dutch version of the MBI (Maslach, & Jackson, 1981b) and the Oldenburg Burnout Inventory (OLBI; Demerouti et al. 2003). The OLBI has only been used in the Flemish sample.

As shown in Table 41, the difference in correlations between the BAT-12 and the BAT-23 on the one hand and both other burnout questionnaires on the other is quite small. For Flanders, the difference for the three subscales of the UBOS varies from .01 to .06 and for both OLBI from .00 to .05. For the Netherlands the difference for the UBOS varies from .00 to .03, depending on the subscale.

As Table 41 also displays, the bold correlations of the BAT-12 total scores with the UBOS and OLBI subscales are (almost) identical to those of the BAT-23. So taken together it can be concluded that the convergent validity of the BAT-12 does not differ from that of the BAT-23.

Table 41: Correlations between burnout scales (BAT, UBOS, OLBI) in Flanders (N = 1,500) and the Netherlands (N=1,500)

		Flanders					Netherlands		
		UBOS			OLBI		UBOS		
BAT-12		**EX**	**CY**	**CO**	**EX**	**DE**	**EX**	**CY**	**CO**
1	*Exhaustion*	.82	.59	-.40	.76	62	.79	.64	-.28
2	*Mental distance*	.65	.80	-.55	.68	.72	.68	.73	-.37
3	*Emotional impairment*	.55	.49	-.34	.47	.47	.64	.63	-.31
4	*Cognitive impairment*	.54	.49	-.45	.53	.49	.67	.63	-.30
5	*Total BAT-12*	**.80**	**.74**	**-.53**	**.74**	**.73**	**.76**	**.72**	**-.35**
BAT-23									
6	*Exhaustion*	.85	.64	-.46	.80	.67	.76	.61	-.26
7	*Mental distance*	.62	.79	-.58	.64	.76	.68	.73	-.38
8	*Emotional impairment*	.60	.53	-.36	.52	.51	.62	.60	-.31
9	*Cognitive impairment*	.57	.51	-.46	.55	.50	.65	.61	-.30
10	*Total BAT-23*	**.80**	**.74**	**-.55**	**.77**	**.74**	**.76**	**.72**	**-.35**

Note: EX = exhaustion, CY = cynicism, CO= competence (personal accomplishment) DE = disengagement

6.2. Work engagement, work addiction and boredom

To investigate the discriminant validity of the BAT with regard to work engagement (UWES; Schaufeli & Bakker, 2004a), work addiction (DUWAS; Schaufeli et al., 2011) and boredom at work (DUBS; Reijseger et al., 2012) the guidelines of Formel and Larcker (1981) were followed (see pp. 56-57).

Discriminant validity exists when the average variance explained (Average Variance Explained, AVE) of a latent factor – in our case burnout, as measured by the BAT-12 – is greater than the squared correlation (R^2) of the relevant latent factors – in our case work engagement (UWES), work addiction (DUWAS) and boredom at work (DUBS).

Tables 42 and 43 show the values of the AVE and R^2 for Flanders and the Netherlands, respectively. As shown in Table 42, the value of AVE for the BAT-12 is higher than the R^2 for the UWES, DUWAS and DUBS, but *lower* than the R^2 for secondary symptoms of the BAT.

Table 42: Average Variance Explained (AVE) and squared correlations (R^2) for work engagement, work addiction, boredom and burnout (Flanders; N = 1,500)

Questionnaire	AVE	R^2				
		UWES	EXC	COM	DUBS	BAT-12
Work engagement (UWES)	.76	--	--	--	--	--
Working excessively (EXC)	.51	.07	--	--	--	--
Working compulsively (COM)	.53	.00	.59	--	--	--
Boredom (DUBS)	.58	.41	.12	.02	--	--
BAT-12	.54	.43	.04	.26	.20	--
BAT secondary symptoms	.52	.17	.07	.23	.03	.62

Note: UWES = Utrecht Work Engagement Scale; EXC = DUWAS Working excessively; COM = DUWAS Working compulsively; DUBS = Dutch Boredom Scale.

Table 43 indicates that also in the Netherlands the value of AVE for the BAT-12 is higher than the R^2 for the UWES, but lower than the R^2 for secondary complaints.

Table 42: Average Variance Explained (AVE) and squared correlations (R^2) for work engagement, work addiction, boredom and burnout (the Netherlands; N = 1,500)

Questionnaire	AVE	R^2	
		UWES	BAT-12
Work engagement (UWES)	.80	--	--
BAT-12	.64	.12	--
BAT secondary symptoms	.62	.07	.72

Note: UWES = Utrecht Work Engagement Scale.

Conclusion: The discriminant validity of the BAT-12 has been demonstrated with regard to the UWES, DUWAS and DUBS, but – as with the full version of the BAT (Tables 18 and 19, p. 57) – *not* with regard to secondary burnout symptoms. This means that scores on the BAT-12 do indeed measure something other than (lack of)work engagement, workaholism (excessive and compulsive working) and boredom at work. The scores on the BAT-12, on the other hand, *cannot* be distinguished from secondary burnout complaints, which is in accordance with the fact that burnout can be regarded as a syndrome, also including secondary symptoms. The same applies, mutatis mutandis, to both dimensions of work addiction; these too cannot be distinguished from each other because they refer to the same underlying concept.

7. Relationships with job demands, job resources, outcomes, and personality

To map out the nomological network of the BAT-12, the relationships were examined with stressors (job demands and work-life conflict), job resources, organizational outcomes and personality factors. In doing so, the Job Demands-Resources (JD-R) model was leading (Schaufeli & Bakker, 2004b; Bakker & Demerouti, 2016) and an online questionnaire based on this model was used: the Energy Compass (Schaufeli, 2015, 2017b). More detailed information about the Energy Compass is included in Appendix 1. In the Dutch sample, a large number of job demands, job resources and organizational outcomes were measured, while the Flemish sample included various personality factors.

Table 44 shows the correlations between the BAT-12 on the one hand and seven job demands and work-life conflict on the other. For comparison, correlations with the total score of the BAT-23 are also included.

Table 44: Correlations (r) of the BAT-12 and BAT-23 with job demands (the Netherlands, N = 1,500)

Job demands	B12-EX	B12-MD	B12-EI	B12-CI	BAT-12	BAT-23
Qualitative demands	.32	.25	.24	.28	.31	.32
Emotional demands	.37	.27	.31	.27	.34	.34
Physical demands	.32	.25	.24	.16	.28	.29
Bureaucracy	.46	.46	.44	.45	**.51**	**.51**
Role conflict	**.54**	**.59**	**.54**	**.53**	**.63**	**.62**
Interpersonal conflict	**.53**	**.60**	**.60**	**.57**	**.65**	**.65**
Undesirable behavior	.44	.49	**.54**	**.53**	**.56**	**.55**
Work-life conflict	**.53**	**.52**	**.54**	**.50**	**.59**	**.58**

Note: All correlations, p <.001; Correlations ≥ .50 are in bold. B12-EX = BAT-12 exhaustion, B12-MD = BAT-12 mental distance, B12-EI = BAT-12 emotional impairment, B12-CI = BAT-12 cognitive impairment, BAT-12 = Total score (12 items), BAT -23 = Total score (23 items).

As can be expected on the basis of the JD-R model (see also, p. 58), all relationships are positive. The strongest relationships are found with role conflict, interpersonal conflict, work-life conflict and socially undesirable behavior (i.e., bullying, intimidation,

discrimination, and sexual harassment)[20]. It is also striking that the exhaustion is at least moderately related (*r* >. 30) with *all* job demands.

Finally, and most importantly, the total scores of the BAT-12 and BAT-23 correlate identically with all job demands; there is only one negligible difference of .01. A comparison of the correlations of the subscales of the BAT-12 in the Table 44 with those of the BAT-23 displayed in Table 20 (p. 59) yields the same picture, with a difference of .03 at most. In short, the BAT-12 and BAT-23 exhibit the same pattern of correlations with job demands.

To determine the unique contribution of each job demand, a number of regression analyzes have been carried out, the results of which are shown in Table 45. Role conflict, interpersonal conflict, socially undesired behavior, and work-life conflict play the most important role[21]. Taken together, job demands explain more than 40% of the variance of the subscale scores and more than 50% of the variance of the both total scores of the BAT.

Table 45: Regression coefficients (β) of the BAT-12 and BAT-23 with job demands (the Netherlands, N = 1,500)

Job demands	B-12-EX	B-12-MD	B-12-EI	B-12-CI	BAT-12	BAT-23
Qualitative demands	--	.06*	--	--	--	--
Emotional demands	.08***	--	--	--	--	--
Physical demands	.13***	.05*	--	--	.05**	.07***
Bureaucracy	.23***	.15***	.18***	.15***	.20***	.20***
Role conflict	--	--	--	.06*	--	--
Interpersonal conflict	.22***	.30***	.14***	.15***	.23***	.24***
Undesirable behavior	12***	.23***	.26***	.19***	.23***	.22**
Work-home conflict	.06*	.11***	.20***	.24***	.17***	.17***
Explained variance	40%	44%	44%	41%	52%	52%

Note: * p <.05, ** p <.01, ** p <.001. B-12-EX= BAT-12 exhaustion, B-12-MD = BAT-12 mental distance, B-12-EI = BAT-12 emotional impairment, B-12-CI = BAT-12 cognitive impairment, BAT-12 = Total score (12 items), BAT-23 = Total score (23 items).

The pattern of the β-coefficients from Table 45 is similar to that in Table 21 (p. 60), which means that the shortened subscales are associated with job demands in the same

[20] Correlations with negative organizational changes, pace of organizational change, and mental work demands are below .30 for all burnout scales and are therefore not included in Table 44.

[21] Please note that the relationship with work-home conflict does not mean that burnout is "caused" by personal problems. This relationship may alternatively be due to employees feeling fatigued and burdened at home due to work overload – i.e. problems at work that negatively affect their life at home and may therefore cause conflict.

way as the corresponding longer versions. The same applies to the total scores for the BAT-12 and BAT-23 (see Table 45). In other words, the unique relationships of specific job demands are virtually the same for the BAT-12 and BAT-23.

Table 46 shows the correlations of the BAT with the main job resources[22]. As can be expected on the basis of the JD-R model, all correlations in this table are negative. Role clarity, meeting expectations and team atmosphere are the most important job resources. It is striking that mental distance is most strongly associated with all job resources. A comparison with Table 22 (p. 61) shows that the relationships of the shortened subscales of the BAT with all job resources is practically identical to that of the corresponding longer versions; differences being .04 at most. The same applies to both total scores, which differ .01 at most (see Table 46).

Table 46: Correlations (r) of the BAT-12 and BAT-23 *with job resources (the Netherlands, N = 1,500)*

Job resources	B12-EX	B12-MD	B12-EI	B12-CI	BAT-12	BAT-23
Role clarity	-.29	**-.36**	**-.32**	**-.31**	**-.36**	**-.35**
Meeting expectations	-.27	**-.33**	**-.31**	**-.30**	**-.34**	**-.33**
Team atmosphere	**-.34**	**-.43**	**-.36**	-.29	**-.40**	**-.40**
Team effectiveness	-.26	**-.34**	-.24	-.22	**-.30**	**-.30**
Person-job fit	-.25	**-.38**	-.19	-.22	-.29	**-.30**
Possibilities for learning	-.18	**-.32**	-.11	-.10	-.20	-.21

Note: All correlations, p <.001; Correlations ≥ .30 are in bold. B12-EX = BAT-12 exhaustion, B12-MD = BAT-12 mental distance, B12-EI = BAT-12 emotional impairment, B12-CI = BAT-12 cognitive impairment, BAT-12 = Total score (12 items), BAT -23 = Total score (23 items).

Table 47 below shows the results of the regression analyzes with various job resources as independent variables. The this table shows that role clarity, team atmosphere and person-job fit make the most important unique contributions to explaining the BAT-12 scores. The pattern of β-values from Table 47 correspond to that of the BAT-23 (see Table 23, p. 62). Furthermore, it appears that the six job resources contribute about equally to explaining the total scores of the BAT-12 and the BAT-23.

[22] The correlations with social support from colleagues and supervisors, task variation, job control, alignment, performance feedback, career possibilities, availability of tools, use of skills, participation in decision-making, fair pay, value congruence with the organization, and engaging leadership are lower than .30 for all burn-out scales and are therefore not included in Table 46.

Table 47: Regression coefficients (β) of the BAT-12 and BAT-23 with job resources (the Netherlands, N = 1,500)

Job resources	B-12-EX	B-12-MD	B-12-EO	B-12-CO	BAT-12	BAT-23
Role clarity	-.13***	-.14***	-.17***	-.18***	-.17***	-.17***
Meeting expectations	--	--	-.14***	-.14***	-.10**	-.08**
Team atmosphere	-.23***	-.27***	-.33***	-.19***	-.30***	-.28***
Team effectiveness	--	--	-.13**	--	-.10**	-.09*
Person-job fit	-.10**	-.20***	-.20***	-.14***	-.14***	-.11***
Possibilities for learning	--	--	-.09*	-.14***	--	--
Explained variance	14%	24%	17%	15%	21%	20%

Note: * p <.05, ** p <.01, ** p <.001. B-12-EX = BAT-12 exhaustion, B-12-MD = BAT-12 mental distance, B-12-EI = BAT-12 emotional impairment, B-12-CI = BAT-12 cognitive impairment, BAT-12 = Total score (12 items), BAT-23 = Total score (23 items).

Table 48 provides an overview of the correlations of the BAT with various organizational outcomes [23] In particular, work-related absenteeism and turnover intention are positively associated with burnout, while in-role performance is negatively associated with burnout. The correlations from Table 48, which refer to the subscales of the BAT-12, hardly differ from the correlations from Table 24 (p. 62), which refer to the BAT-23; differences being .02 at most. Finally, the correlations of the outcomes with both total scores of the BAT are almost identical. Again, the BAT-12 and BAT-23 behave practically the same, also with regard to eight different organizational outcomes.

Table 48: Correlations (r) of the BAT-12 and BAT-23 with organizational outcomes (the Netherlands, N = 1,500)

Organizational outcomes	B-12-EX	B-12-MD	B-12-EO	B-12-CO	BAT-12	BAT-23
Organizational commitment	-.23	**-.37**	-.18	-.16	-.27	-.27
Team commitment	-.26	**-.37**	-.23	-.19	**-.30**	**-.30**
Job satisfaction	**-.35**	**-.49**	-.29	-.28	**-.40**	**-.40**
Turnover intention	**.42**	**.50**	**.38**	**.42**	**.49**	**.49**
Workability	**-.32**	-.28	**-.34**	-.26	**-.34**	**-.34**
Work related absenteeism	**.36**	**.34**	**.34**	**.38**	**.38**	**.37**
In-role performance	-.25	**-.34**	**-.37**	**-.32**	**-.36**	**-.35**
Job performance	-.25	**-.30**	-.29	-.29	**-.32**	**-.31**

Note: All correlations, p <.001; Correlations ≥ .30 are in bold. B12-EX = BAT-12 exhaustion, B12-MD = BAT-12 mental distance, B12-EI = BAT-12 emotional impairment, B12-CI = BAT-12 cognitive impairment, BAT-12 = Total score (12 items), BAT -23 = Total score (23 items).

[23] The correlations with duration and frequency of absenteeism, subjective health, and extra-role performance are lower than .30 for all burnout scales and are therefore not included in the Table 48.

Table 49 shows the results of the regression analyzes that were carried out with regard to the organizational outcomes. This shows that turnover intention, workability and work-related absenteeism make the most important unique contributions to explaining the BAT-12 scores. The pattern of β-values from Table 49 corresponds to that of the BAT-23 (see Table 25, p. 64). Finally, it appears that all organizational outcomes contribute in an almost identical way to explaining the variances of the total scores of the BAT-12 and the BAT-23.

Table 49: Regression coefficients (β) of the BAT12 and BAT-23 with organizational outcomes (the Netherlands, N = 1,500)

Organizational outcomes	B-12-EX	B-12-MD	B-12-EO	B-12-CO	BAT-12	BAT-23
Organizational commitment	--	-.12**	.16***	--	--	--
Team commitment	.09*	--	--	--	--	--
Job satisfaction	--	-.21***	--	--	--	--
Turnover intention	.35***	.40***	.33***	.40***	.44***	.44***
Workability	-.14***	--	-.11*	-.09*	-.10**	-.10**
Work related absenteeism	.19***	.12**	.17***	.11**	.17***	.17***
In-role performance	--	--	-.20***	-.16**	-.11**	-.10*
Job performance	--	-.08*	-.12**	-.13**	-.10*	-.09**
Explained variance	28%	43%	31%	28%	38%	37%

Note: * p <.05, ** p <.01, ** p <.001. B-12-UI = BAT-12 exhaustion, B-12-MD = BAT-12 mental distance, B-12-EI = BAT-12 emotional impairment, B-12-CI = BAT-12 cognitive impairment, BAT-12 = Total score (12 items), BAT-23 = Total score (23 items).

Table 50 shows the correlations of the BAT with a number of personality characteristics[24].

Table 50: Correlations (r) of the BAT-12 and BAT *with personality (Flanders N = 1,500)*

Personality characteristics	B-12-EX	B-12-MD	B-12-EI	B-12-CI	BAT-12	BAT-23
Neuroticism	**.45**	**.40**	**.41**	**.38**	**.50**	**.51**
Consciousness	-.18	-.21	-.13	**-.37**	-.26	-.28
Self-efficacy	-.21	-.27	-.21	-.29	**-.30**	**-.30**
Optimism	**-.48**	**-.56**	**-.39**	**-.39**	**-.56**	**-.57**
Resilience	-.28	**-.32**	-.29	**-.34**	**-.37**	**-.38**

Note: All correlations, p <.001; Correlations ≥ .30 are in bold. B12-EX = BAT-12 exhaustion, B12-MD = BAT-12 mental distance, B12-EI = BAT-12 emotional impairment, B12-CI = BAT-12 cognitive impairment, BAT-12 = Total score (12 items), BAT -23 = Total score (23 items).

Neuroticism and optimism show the strongest relationships with burnout. The pattern and size of the correlations in Table 50 are very similar to those in Table 26 (p. 64), the

[24] Correlations with extroversion, openness, agreeableness, flexibility, and setting one's own boundaries are lower than .30 for all burnout scales and are therefore not included in Table 49.

difference being .04 at most. Furthermore, the correlations of the personality characteristics with both versions of the BAT are almost identical.

Table 51 provides an overview of the results of the regression analyzes in which personality characteristics act as independent variables. As shown in this table, optimism is most strongly and negatively associated with burnout, followed by neuroticism, which is positively related. The pattern of β-coefficients in Table 51 is very similar to that in Table 27 (p. 65). The four personality characteristics also contribute in an almost identical way to explaining the variance in the total scores of the BAT-12 and the BAT-23.

Table 51: Regression coefficients (β) from the BAT12 and BAT-23 with personality (Flanders, N = 1,500)

Personality characteristics	B-12-EX	B-12-MD	B-12-EI	B-12-CI	BAT-12	BAT-23
Neuroticism	. 27***	.12***	.29***	.19***	.27***	.28***
Consciousness	--	-.07**	--	-.26***	-.12***	-.13***
Self-efficacy	-.34***	-.48***	-.24***	-.19***	-.39***	-.39***
Optimism	--	--	--	-.06*	--	--
Explained variance	29%	33%	21%	26%	39%	40%

Note: * p <.05, ** p <.01, ** p <.001. B-12-EX = BAT-12 exhaustion, B-12-MD = BAT-12 mental distance, B-12-EI = BAT-12 emotional impairment, B-12-CI = BAT-12 cognitive impairment, BAT-12 = Total score (12 items), BAT-23 = Total score (23 items).

In conclusion: the observed relationships between job demands and job resources on the one hand and burnout, as measured by the BAT-12, on the other hand agree with the predictions of the JD-R model. This means that job demands are positive and job resources are negatively associated with the BAT-12. Regarding job demands, the strongest positive correlations are found with role conflict, interpersonal conflict, work-life conflict, and socially undesirable behavior. As far as job resources are concerned the strongest negative correlations are found with role clarity, meeting expectations, team atmosphere, and person-job fit. Also according to the expectations of the JD-R model, job demands are more strongly associated with the BAT-12 than job resources; 52% of the variance is explained by job demands against only 21% is by job resources.

The main finding, however, is that scores on the BAT-12 and on the BAT-23 are almost identically related to job demands and job resources; the differences in correlations are .03 at most. This also applies – mutatis mutandis – to organizational outcomes and

personality characteristics, these are also similarly related to the BAT-12 as to the BAT-23. The main organizational outcomes are turnover intention and work-related absenteeism, both positively associated with the BAT-12, and workability, which is negatively associated with the BAT-12. Neuroticism (negative) and optimism (positive) are the main personality characteristics that are associated with the BAT-12.

In short, for studying the relationships with job demands, job resources, organizational outcomes and personality characteristics, the shortened BAT-12 can be used just as well as the original long version of the BAT which has 23 items.

8. The general version of the BAT-12

A shortened, *general* version of the BAT has also been constructed. This questionnaire consists of the corresponding items of the work-related version, but without reference to "work" in all items except those of mental distance subscale.

Table 52: The short general version of the BAT

Exhaustion
1. I feel mentally exhausted
2. I find it hard to recover my energy
3. I feel physically exhausted
Mental distance
4. I struggle to find any enthusiasm for my work
5. I feel a strong aversion towards my job
6. I'm cynical about what my work means to others
Emotional impairment
7. I feel unable to control my emotions
8. I do not recognize myself in the way I react emotionally
9. I may overreact unintentionally
Cognitive impairment
10. I have trouble staying focused
11. I have trouble concentrating
12. I make mistakes because I have my mind on other things

For psychometric evaluation, a Flemish online sample was used, the characteristics of which are shown in Table 28 (p. 66). This sample consists of both employed employees (N = 416; 81%) and employees who are on sick leave (N = 96; 19%).

8.1. Reliability

Table 53 displays the means, standard deviations, and internal consistencies of the general version of the BAT-12 for those who are currently working and on sick leave, respectively.

The internal consistency of the shortened general BAT-12 is by definition less than that of the original longer version (see Table 29, p. 67), but still more than sufficient in both groups.

Table 53: Means (M), standard deviations (SD) and internal consistency (coefficient α) of the general version of the BAT-12 in Flanders.

BAT-12 general version	# items	Employed (N = 416)			Sick listed (N = 96)		
		M	SD	α	M	SD	α
Exhaustion	3	3.55	.93	.91	4.33	.73	.92
*Mental distance**	3	2.92	1.07	.88	3.54	.95	.78
Emotional impairment	3	3.14	.96	.94	3.99	.99	.94
Cognitive impairment	3	3.19	.81	.92	3.84	.96	.94
Total BAT-12	12	3.20	.79	.94	3.95	.67	.90
*Secondary symptoms**	10	3.34	.74	.94	3.93	.68	.90

Note: * Identical to the work-related version of the BAT.

Not surprisingly, those on sick leave score significantly higher on all general BAT-12 scales than those at work: exhaustion ($t_{(510)} = 8.95$; $p < .001$), mental distance ($t_{(510)} = 5.30$; $p < .001$), emotional impairment ($t_{(510)} = 7.87$; $p < .001$), cognitive impairment ($t_{(510)} = 6.76$; $p < .001$), BAT core symptoms ($t_{(510)} = 9.01$; $p < .001$) and BAT secondary symptoms ($t_{(510)} = 16.67$; $p < .001$). This means that those who are on sick leave are more likely to suffer from burnout complaints than those who are not on sick leave. This is also to be expected given the fact that 76% indicate that their absence is related to work and that 63% have been treated for burnout in the past 5 years.

In a sample of 964 Flemish respondents who were approached online, the general version of the BAT was administered three times, with a time interval of 6 months (for more details about this sample, see 5.2, p. 49). As can be seen from Table 54, the

stability coefficients of the shortened general version hardly differ from those of the longer general version (Table 30, p. 67).

The stability coefficients of the general BAT-12 subscales range from .65 to .73 over a 6 month period, and from .59 to .68 over a one year period (Table 54). Furthermore, it appears that the test-retest reliability of the total BAT12 (core symptoms) is higher than that of its subscales.

Table 54: Test-retest reliability of the general version of the BAT-12 (stability-coefficient r_t)

Scale	T1 -T2 (N = 597)	T2-T3 (N = 368)	T1-T3 (N=447)
Exhaustion	.72	.73	.68
Mental distance	.68	.65	.59
Emotional impairment	.68	.66	.63
Cognitive impairment	.69	.70	.62
Total BAT-12	.79	.78	.71

8.2. Correlations between the general and work related versions

The general shortened version of the BAT-12 is almost perfectly correlated with the longer 23-item-item version: exhaustion (*r* = .96), mental distance (*r* = .97), emotional impairment (*r* = .98), cognitive impairment (*r* = .99), and total score (*r* = .99). Such correlations suggest that the original and short general versions of the BAT are identical.

It appears that the correlations between the work related BAT-12 and the general version in the working group are also very high: exhaustion (*r* = .78), emotional impairment (*r* = .81), cognitive impairment (*r* = .79), and BAT -12 total score (*r* = .93). This means that the corresponding subscales overlap between 60-65%, while the common variance of both total scores is 86%. Please note that the shortened mental distance scale is identical for both versions of the BAT-12.

8.3. Factor structure

An exploratory factor analysis with all work and non-work related items yielded a 4-factor model in which both types of items formed a separate factor for each subscale: exhaustion (21% explained variance, factor loadings vary from .63 to .79), emotional impairment (23% explained variance, factor loadings vary from .75 to .81), cognitive impairment (21% explained variance, factor loadings vary from .66 to .79), and mental distance (14% explained variance, factor loadings vary from .73 to .75). Furthermore, the first unrotated factor explains 58% of the variance and the loadings on this factor vary from .67 to .83. This suggests that the general BAT-12 measures one underlying concept, that consists of four very closely related aspects.

The very high correlations between the two shortened versions of the BAT-12 and the fact that the items of both versions form joint rather than different factors indicates that the general version actually measures the same as the work-rated version of the BAT-12.

A confirmatory factor analysis was then performed on the items of the general version of the BAT-12, testing three models: (1) a 1-factor model that assumes that all items load on one single factor (Model 1); (2) a 4-factor model with correlated factors: exhaustion, mental distance, and emotional and cognitive impairment (Model 2); and (3) a second-order model in which all four factors all load on one underlying latent factor (Model 3). This model is consistent with the idea that burnout is a syndrome consisting of four closely related symptoms that refer to one underlying condition. The three models were tested simultaneously in the employed and in the sick listed group (multiple group method). Table 55 shows the fit indices of the three models.

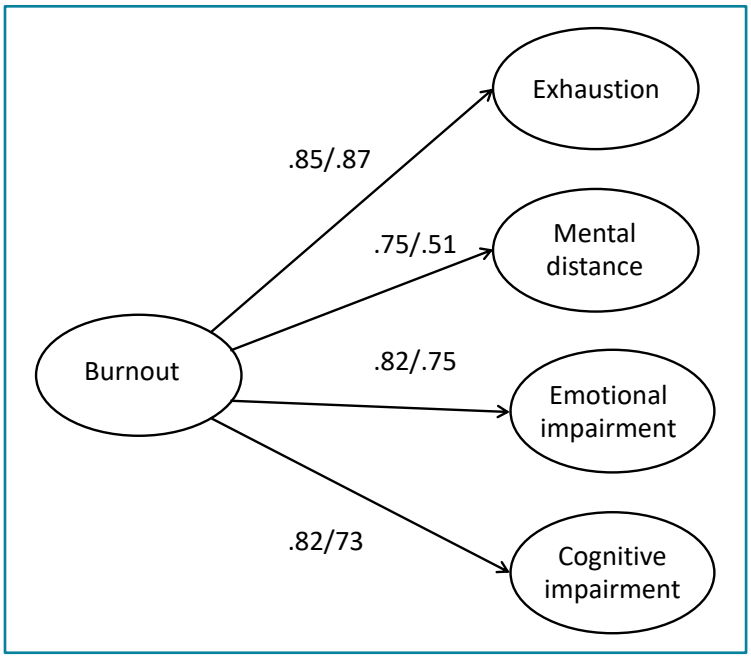
Table 55: Confirmative factor analysis of the general version of the BAT-12 (employed N = 416; sick listed N = 96)

Model		χ^2	df	CFI	TLI	RMSEA
1	*1-factor model*	1845.23	108	.68	.61	.18
2	*4-factor model*	191.49	96	.98	.98	.04
3	*Second-order model*	193.43	100	.98	.98	.04

Note: χ^2 = chi-square; df = degrees of freedom; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation.

The 1-factor model (Model 1) does not fit to the data, but both the 4-factor model (Model 2) and the second-order model (Model 3) fit the data well. The fit of Model 2 and Model 3 does not differ significantly ($\Delta\chi^2 = 1.94$, df = 4, n.s.). However, the second-order model is preferred because it corresponds to the basic idea underlying the BAT, namely that burnout is a syndrome. Figure 8 shows the second-order factor charges.

Figure 8: The structure of the general version of the BAT-12



Note: employed/sick listed.

As with the longer general version of the BAT-23 (see Figure 6, p. 69), the second-order factor loadings are somewhat lower in the group on sick leave. This is especially true for mental distance, probably because it doesn't play such a big role among those on sick leave compared to the other three aspects of burnout. This is unsurprising because since they do not work this is not top of mind.

8.4. Validity

To investigate whether those on sick leave also score – as can be expected – unfavorably on other indicators of well-being, correlations between the general version of the BAT-12 and five different indicators of well-being are computed. These indicators are:

- Job satisfaction: "To what extent are you, overall, satisfied with your job?" Scoring: 1 = very dissatisfied, 2 = dissatisfied, 3 = not dissatisfied / not satisfied, 4 = satisfied, 5 = very satisfied.
- Subjective health: "In general, what do you think of your health?" Scoring: 1 = poor, 2 = moderate, 3 = good, 4 = very good, 5 = excellent.
- Happiness: "On a scale from 0-10, how happy do you feel?"
- Feeling burned out: "Do you feel burned-out because of your job?". Scoring: 1 = never, 2 = rarely, 3 = occasionally, 4 = often, 5 = always.
- Depression: Beck Depression Inventory II (Beck, Steer & Gregory, 1996; 21 items, α = .88).

Table 56: Correlations of the general version of the BAT-12 with indicators of well-being (N = 96)

BAT-12 general version	Job satisfaction	Subjective health	Happiness	Feeling burned-out	Depression (BDI)
Exhaustion	-.37**	-.46**	-.50**	.66**	.50**
Mental distance	-.62**	-.01	-.30**	.39**	.31**
Emotional impairment	-.13	-.15	-.46**	.44**	.48**
Cognitive impairment	-.27**	-.25**	-48**	.44**	.37**
Total BAT-12	-.43**	-.30**	-.56**	.62**	.53**
Secondary symptoms	-.31**	-.32**	-.46**	.54**	.51**

Note: ** $p < .01$.

All correlations in Table 56 are in the expected direction, although three are not significant. Exhaustion correlates moderately to strongly with all indicators. Mental distance is mainly related to job satisfaction, but not subjective health. Emotional impairment is most strongly associated with depression (positive) and happiness (negative), while cognitive impairment is related to happiness (negative) and to feeling burned-out (positive). The overall score on the BAT is moderately related to all wellbeing indicators, as well as secondary burnout symptoms. The fact that subjective health is least strongly related to the general BAT12 might be explained by the fact that this may particularly refer to physical health. The strongest relationships are found with the indicators of mental health (i.e., depression and feeling burned-out).

Overall, the correlations displayed in Table 56 confirm the validity of the general version of the BAT-12. This means that those who score high are not satisfied with their

work and do not feel very healthy and happy, but instead feel burned-out and depressed.

9. Conclusions

The original 23-item version of the BAT, which measures the four core symptoms of burnout, has been shortened to 12 items based on a combination of Rasch and content analysis. Like the BAT-23, the BAT-12 consists of four subscales (exhaustion, mental distance, emotional and cognitive impairment), each containing three items.

The fact that the BAT-12 meets the requirements of the Rasch model means that the items are of good psychometric quality. In addition, Rasch analyses suggests that the BAT-12 is a unidimensional scale, meaning that all items measure the underlying burnout construct. In other words, the total score on the BAT-12 indicates the degree of burnout.

This is further confirmed by the fact that the second-order factor model, which assumes that the four subscales of the BAT-12 load on one common, underlying factor, fits the data well. This second-order model is consistent with the view that burnout can be understood as a syndrome consisting of four types of strongly related symptoms that refer to a common mental state – i.e., burnout. At the same time, however, a model with four strongly correlated factors also fits well with the data. As with the BAT-23, this leads to the conclusion that the subscales of the BAT-12 can also be considered separately. In other words, although four core aspects of burnout can be distinguished, they are so closely interrelated that one can speak of a single burnout construct.

The overall score of the BAT-12 correlates almost perfectly with that of the BAT-23, while more or less the same applies to the corresponding subscales. In short, the overlap between the two versions is so strong that it can be assumed that both are identical and measure the same burnout construct.

The internal consistency of the BAT-12 is very good ($\alpha > .85$) but somewhat lower than that of the BAT-23, because coefficient α decreases by definition with the number of items. However, the difference is very small and amounts to only .04 for the total score

in Flanders and .02 for the Netherlands. The test-retest reliability of the BAT-12 over a 6 and 12 month period is also good and is – with a difference of only .02 – almost identical to that of the BAT-23.

The convergent validity of the BAT-12 was investigated by examining its relationship with two other commonly used burnout questionnaires (MBI and OLBI). It appears that the convergent validity of the BAT-12 does not differ from that of the BAT-23. That is, the correlations with the MBI and the OLBI are practically the same with a difference of .06 at most. In particular, both versions of the BAT are strongly related to the exhaustion and mental distance subscales of both other burnout instruments. Their common variance ranges from 50% -65%.

The divergent validity of the BAT-12 was established with regard to questionnaires that measure work engagement, workaholism, and boredom at work. It was established – as with the BAT-23 – that the BAT-12 does indeed measure a different concept. This does not apply to secondary burnout symptoms, which cannot be distinguished from the core symptoms of burnout, as measured by the BAT-12. This is not surprising because the core and secondary symptoms both relate to burnout.

The content validity of the BAT-12 was investigated by analyzing its relationships with a large number of job demands, job resources, personality characteristics, organizational outcomes, and personality factors. The JD-R model was used as a conceptual framework for establishing content validity. More specifically, the following was observed with regard to the BAT-12:

- A *positive* relationship with job stressors, such as work-life conflict, interpersonal conflict, role conflict, bureaucracy, and socially undesirable behavior.
- A *negative* relationship with job resources, such as team atmosphere, role clarity, and person-job fit.
- A *negative* relationship with job satisfaction, turnover intention, and job performance and a *positive* correlation with absenteeism and workability.
- A *positive* relationship with neuroticism and a *negative* relationship with optimism, conscientiousness, resilience and self-efficacy.

The overall conclusion is that the shortened version of the work-related BAT can, psychometrically speaking, be used just as well as the original longer version.

The same online sample was used for psychometric research with the shortened general version of the BAT as for the full general version. This sample consisted of 416 workers, who also completed the (shortened) work-related version, and 96 employees who are on sick leave, who only filled in the general version. The main findings are:

- The shortened general BAT scales have a very good internal consistency and the test-retest reliability (across 6 and 12 months) is also good and hardly deviates from that of the longer, general version.
- The correlations between the short work-related and general scales of the BAT are very high. Furthermore, the work-related and general items constitute a separate factor for each subscale. Both findings show that the items of both shortened versions overlap so much that it can be assumed that they measure the same thing.
- Confirmatory factor analyzes show that a second-order model, in which burnout is interpreted as a syndrome of four closely related factors that refer to one underlying condition, fits well with the data. The fact that all items load high on the first unrotated factor of the exploratory factor analysis also points in this direction. In practical terms, this means that in addition to the four subscale scores, the total score of the general short BAT can also be used.
- The validity of the general version of the BAT is supported by the moderate to strong correlations with a number of wellbeing indicators, such as job satisfaction, subjective health, happiness, feeling burned-out, and depression.

The conclusion is that the shortened version of the general BAT can be used as well as the original, longer version.

PART IV: PRACTICAL USE

1. Administration and scoring

Filling out the entire BAT (see Appendix 2) takes about 5 minutes – and even less for the short version – and can be done either individually as part of a burnout assessment procedure or collectively. In the latter case the BAT can used, for instance, as a screening or benchmarking tool in organizations. The general, context-free version of the BAT can be used among those who are currently not working, for instance, for monitoring their progress in treatment, counseling or return-to-work programs.

Because the term "burnout" might have a negative connotation, it is not mentioned in the title of the questionnaire. Instead, the neutral "work experience survey" (or "well-being survey" for the general version) is used.

The test instruction reads[25]: "*The following statements are about how you experience your work and how you feel about it. Please indicate how often each statement applies to you by ticking the most appropriate answer.*"

The average scores on the BAT scales are calculated by adding the scores on all items of a particular subscale and then divide this sum by the number of items of that scale. The following scoring categories are used: 1 "never", 2 "rarely", 3 "sometimes", 4 "often", and 5 "always". This means that the value of the average scale scores varies from 1 to 5. By adding the scores on *all* BAT items and dividing the sum by 23 (or by 12 for the short version), the total score is obtained, which also ranges between 1 and 5. The same procedure is followed for calculating the total score of the 10 secondary symptoms.

The total score of the (shortened) BAT can be used to assess the level of burnout, whereas the scores on the four core dimensions, supplemented with secondary

[25] *For the general version, the instruction reads: "The following statements are about how you feel. Please indicate how often each statement applies to you by ticking the most appropriate answer".*

symptoms, may further differentiate the picture. This kind of differentiation is particularly important for individual burnout assessment.

2. Procedure

Statistical norms and clinical cut-off values have been calculated for the BAT. Those are based on a representative – and in principle healthy – sample and a group that suffers from severe burnout, respectively. For the time being, only clinical cut-offs for Flanders are available; the next version of the manual will also include Dutch cut-offs.

Using statistical norms, BAT scores can be compared with the national workforce, that serves as a reference group. This is important if one is interested in the relative level of burnout of an individual or group, as compared to the average Flemish or Dutch employee. Using cut-off scores, burnout "cases" can be distinguished from "non-cases". This is imperative for individual burnout assessment and for screening organizational members who are at risk for burnout.

2.1. Statistical norms based on a representative sample

For norming the BAT by using a representative sample of the Dutch and Flemish workforce, a classification was made into four categories: "low", "average", "high", and "very high". Considerations regarding the standard measurement error and the distribution of the scores played a role in this decision. The aim was to strike a balance between practical relevance and a statistically meaningful distinction. Table 56 shows the lower and upper boundaries of the four burnout levels.

Table 57: BAT scoring categories

Burnout level	Lower boundary		Upper boundary
Very high	95th percentile	≥ score	
High	75th percentile	≤ score	< 95th percentile
Average	25th percentile	≤ score	< 75th percentile
Low		score	< 25th percentile

This way, it is possible to assess the level of burnout of individuals and groups, based on a comparison with the "average" Flemish or Dutch employee. This may then lead to statements such as: "This group (or person) has a high (or low) level of burnout compared to the average Flemish (or Dutch) employee."

2.2. Clinical cut-off scores based on a group with severe burnout

To determine appropriate cut-off scores in Flanders, a group of patients was used that participated in a study on the epigenetics of burnout and depression at KU Leuven (N = 40), supplemented by a group from a burnout prevention and treatment center (N = 27). The participants from both groups (N = 67) were selected on the basis of a clinical interview in which the inclusion and exclusion criteria from Table 58 were used.

Table 58: Selection criteria for severe burnout

Inclusion criteria

A. Physical and mental exhaustion (even after minimal effort), which is the central feature of severe burnout.
B. Reduced mental energy, which is expressed by a lack of initiative and the inability for perseverance (i.e., the ability to keep doing something in spite of obstacles), poor endurance (i.e., the ability to sustain a prolonged stressful effort or activity), and inability to recover.
C. Loss of control or ineffectiveness of the usual coping strategies.
D. Loss of roles (work or social roles) for at least 50%.
E. Distress symptoms (daily) in the following areas:
    a. Cognitive (lack of concentration, memory deficits, inefficient thinking);
    b. Emotional (irritability, instability).
F. Supplemented by (daily) distress symptoms in at least one of the following areas:
    a. Sleep;
    b. Psychosomatic (e.g. chest pain, palpitations, muscle aches, gastro-intestinal problems);
    c. Psychological (e.g. dizziness, oversensitivity to stimuli like sounds, tensions, nervousness, worrying, inability to relax, gloomy mood).
G. These symptoms should be present for at least 6 months.

Exclusion criteria

A. Psychiatric disorder such as mood disorder, anxiety disorder, PTSD, chronic fatigue syndrome, or fibromyalgia.
B. The symptoms mentioned in the inclusion criteria should not be due to substance abuse, medication or a somatic disease (e.g., diabetes, hyperthyroidism or arthritis). Patients who have a somatic disease that does not influence their present symptoms can be included if they fit the inclusion criteria.

To be included in the group with serious burnout, criteria A to F, and H had to be met. Whether criterion G was met could sometimes not be checked because many participants found this question very difficult to answer.

The criteria from Table 58 are derived from the official Dutch guidelines for assessing, preventing and treating work-related mental problems[26] (Van der Klink & Van Dijk, 2014) as well as the diagnostic criteria for "exhaustion disorder" used by the Swedish National Board of Health and Welfare (Glise, Hadzibajramovic, Jonsdottir & Ahlborg, 2010). Both refer to the criteria for neurasthenia, as included in the ICD-10 classification system (Schaufeli et al., 2001[27]. In fact, of all official diagnoses, work-related neurasthenia serves as the best proxy for severe burnout.

This group has been supplemented by participants from an online survey that was conducted in collaboration with *VZW Burn-out.Vlaanderen*. In this study, 215 participants were asked about the criteria from Table 58, as well as about whether or not they had received an official burn-out diagnosis or had been treated for burn-out. Those who met the criteria from Table 58, were included in the group with severe burnout. This was the case for 91 persons, so that in total 158 persons with severe burnout are available for analysis. Table 59 shows that women are overrepresented in this group and that about one-third has been treated for burnout in the past five years. Tellingly, the average work experience is quite high.

[26] See: https://www.nvab-online.nl/sites/default/files/bestanden-webpaginas/MDRL_Overspanning-Burnout.pdf (in Dutch).
[27] In the most recent ICD-11burnout is not considered a medical diagnosis but an "occupational phenomenon".

Table 59: Overview of the group with severe burnout in Flanders (N = 158)

Gender		
Male	25%	
Female	74%	
Average age in yrs. (SD)	42.9 (10.1)	
Job tenure in yrs. (SD)	20.3 (10.6)	
Treated for burnout	35%	
BAT-23 -scales	**M**	**SD**
Exhaustion	3.79	.58
Mental distance	3.07	.85
Emotional impairment	3.01	.82
Cognitive impairment	3.35	.69
Total score	3.37	.53
Secondary symptoms	3.36	.50

Compared to those hose who were selected on the basis of a clinical interview, the level of burnout of those selected through the online survey was either equally high (for exhaustion, $t_{(156)}$ = -.21, n.s.; cognitive impairment, $t_{(156)}$ = 1.37, n.s.; secondary symptoms, $t_{(156)}$ = .80, ns; and the total score, $t_{(156)}$ = 1.86, n.s.) or significantly higher (for mental distance, $t_{(156)}$ = 2.31, $p < .05$ and emotional impairment, $t_{(156)}$ = 2.14, $p < .05$). Therefore, it can be assumed that "real" burnout cases have been selected from the online survey. For that reason, both groups were pooled into one Flemish group of people with severe burnout.

The group with severe burnout was used to calculate clinical cut-off scores for Flanders. A Relative Operating Characteristics Analysis, or ROC analysis (Streiner & Cairney, 2007) was used for this purpose. This method, which evaluates the discriminant power of a questionnaire, simulates for each possible score how well it distinguishes between a "problematic" and a "non-problematic" group. An external criterion should be used to distinguish both groups. In our case we used a "healthy" group of employees, consisting of the representative sample of Flemish workforce[28].

Using the ROC analysis, an optimum cut-off value can be calculated to discriminate "cases" from "non-cases", taking into account both specificity (the probability of a

[28] In order to have a healthy group without severe burnout, those who have been treated for burnout complaints in the past 5 years have been removed. In Flanders, 97 employees reported that they had received treatment for burnout(6.5%), which means that 1403 "healthy" employees remained for the ROC analyzes.

negative result, meaning that someone is rightly not identified as a burnout case – *true negative rate*) and sensitivity (the probability of a positive result, meaning that someone is rightly identified as a burnout case – *true positive rate*). When an optimum ratio is applied, as many burnout cases as possible are correctly classified and thus there is a high sensitivity, while at the same time also as many non-burnout cases as possible are correctly classified as non-burned-out, which reflects a high specificity. The trade-off between specificity and sensitivity can be seen from the ROC curve, in which the number of "correct positives" (sensitivity) is plotted on the Y-axis against the number of "false positives"(1 - the specificity) on the X-axis, and this for all possible values (see Figures 9 and 10).

So, when determining cut-off values, a trade-off must be made between detecting sufficient burnout cases, but at the same time not overestimating the problem (false positives). In fact, we calculated *two* different kinds of cut-off values that allows to distinguish between three groups, using the so-called traffic light model (Notelaers, De Witte, & Van Veldhoven, 2005): (1) a green, "safe" group, (2) an orange, "group at risk" and (3) a red, "burned-out" group.

A BAT score is considered "at risk" when it combines high sensitivity with high specificity. In that case, the cut-off is the value at which the sum of specificity and sensitivity is maximum (Habibazadeh, Habibazadeh, & Yadollahie, 2016). When determining the cut-off score for the burned-out group, a stricter criterion for specificity is used, which must be greater than 0.90 (Baillien, De Witte, Notelaers & Neyens, 2008). This implies that the likelihood for obtaining a false positive result is less than 10%.

Before cut-off values can be determined, the power of the BAT to discriminate between the healthy group and the group with severe burnout is examined. If the Area Under Curve (AUC) is low, the power is weak so that no good distinction can be made between both groups. An AUC value of 0.5 indicates that instead of using the BAT, a coin can just as well be flipped to assign a person to one of the groups, whereas a value of 1 signifies perfect discrimination between the group with and without severe burnout. For our purpose we consider an AUC value higher than 0.60 as sufficient to discriminate between "healthy" and "burned-out" employees; meaning that in that case it is allowed to calculate cut-off scores (Fischer, Bachmann & Jaeschke, 2003).

2.3. Clinical cut-off scores for the general version of the BAT

To calculate cut-off values for the context-free, general version of the original and shortened BAT, a group of healthy employees (N = 756) and part of the group with severe burnout (N = 130) was used for Flanders. The group of "healthy" employees was surveyed on line with the help of various organizations. The group with severe burnout consists of the same participants as the group used for the work-related version of the BAT, but it is not the same size because only a part of them (namely 130 out of 158) also filled out the general version of the BAT. As shown in Tables 29 (p. 67)and 53 (p. 94), the latter group scores significantly higher than the former group on all (shortehed) BAT scales.

Currently not enough data are available of a Dutch group with severe burnout which means that to date only for Flanders clinically validated cut-off values are available for the original and shortened versions of the BAT.

3. Norms and cut-off values for the BAT-23

Both statistical norms and cut-off values are presented below. The former are based on representative samples of the Flemish and Dutch workforce, whereas the latter are based on a group of employees with who suffer from burnout.

3.1. Statistical norms for the employed

Tables 60 and 62 provide an overview of the statistical norms for Flemish employees. A comparison with the average BAT scores of the group from Table 59 (p. 106) with severe burnout reveals that the scores either fall in the "very high"(exhaustion and total-score) or "high" (mental distance, emotional and cognitive impairment) category. This means that those who score "high" are comparable to the group with severe burnout. The fact that the "orange" and "red" cut-off values for the BAT total-score from Table 66 (p. 113; 2.59 and 3.02, respectively) also fall into the "high" category confirms this picture. The entire "very high" category consists of employees who most likely suffer from burnout because the "red" cut-off score falls below the "very high" criterion of Table 60.

Table 60: Norms for the core symptoms of the BAT for Flanders (N = 1,500)

BAT-23	Total	EX	MD	EI	CI
Low	≤. 1.60	≤. 1.75	≤. 1.20	≤. 1.20	≤. 1.80
Average	1.61 – 2.40	1.76 – 2.70	1.21 – 2.40	1.21 – 2.19	1.81 – 2.59
High	2.41– 3.29	2.71 – 3.74	2.41 – 3.59	2.20 – 3.19	2.60 –3.39
Very high	≥ 3.30	≥ 3.75	≥ 3.60	≥ 3.20	≥ 3.40
M	2.08	2.27	1.98	1.82	2.13
SD	.64	.77	.84	.71	.69
SE	.02	.02	.02	.02	.02
Range	1.00-5.00	1.00-5.00	1.00-5.00	1.00-5.00	1.00-5.00

Note: Total = BAT total-score, EX = BAT exhaustion, MD = BAT mental distance, EI = BAT emotional impairment, CI = BAT cognitive impairment, M = mean, SD = standard deviation, SE = standard error of the mean.

The observed percentages of respondents from the norm group who score "high" and "very high" according to the norms of Table 60 may, in fact, deviate from the expected percentages. Table 61 provides an overview of these observed percentages from which it appears that the expected percentage of 5% for "very high" scores is approached quite well.

Table 61: Observed percentages "high" and "very high" on the BAT for Flanders (N = 1,500)

BAT-23	Total	EX	MD	EI	CI
% "High"	20.4	21.5	18.8	24.3	21.1
% "Very high"	5.1	5.5	5.8	5.5	5.2
% Total	25.6	27.0	24.6	29.8	26.3

Note: Total = BAT total-score, EX = BAT exhaustion, MD = BAT mental distance, EI = BAT emotional impairment, CI = BAT cognitive impairment.

The percentage "high" scorers should, together with those scoring "very high", amount to 25%. However, this expected percentage is somewhat less well approached, especially for emotional impairment. This means that someone who scores "high" on emotional impairment does not belong to the highest 25% but the highest 30% of the Flemish workforce.

The average score of the group that suffers from burnout (see Table 59, p.106) falls into the "high" category, as do the "orange" and "red " cut-off values of 2.77 and 3.23 from Table 62 below, respectively. In other words, someone who scores "high" on distress usually experiences severe burnout symptoms, whereas those who score "very high" are most likely suffering from burnout.

Table 62: Norms for the secondary symptoms of BAT for Flanders (N = 1,500)

	Sec	DP
Low	≤. 1.70	≤. 1.00
Average	1.71 – 2.75	1.01 – 1.85
High	2.76 – 3.50	1.86 – 3.14
Very high	≥ 3.51	≥ 3.15
M	2.25	1.55
SD	.72	.74
SE	.02	.02
Range	1.00-5.00	1.00-5.00

Note: Sec = BAT secondary distress symptoms, DP = 4DKL depression, M = mean, SD = standard deviation; SE = standard error of the mean.

As shown in Table 63, the observed percentages of those who score "(very) high" correspond to the expected percentages.

Table 63: Observed percentage "high" and "very high" on secondary symptoms for Flanders (N = 1,500)

	Sec	DP
% "High"	20.5	19.5
% "Very high"	4.8	5.2
% Total	25.3	24.7

Note: Sec = BAT secondary distress symptoms, DP = 4DKL depression, SD = standard deviation

The statistical norm scores for the Netherlands are shown in Table 64. The average BAT scores on most scales are significantly higher than in the Flemish norm group: total-score ($t_{(2998)} = 5.03$, $p <. 001$), exhaustion ($t_{(2998)} = 3.79$, $p < .001$), mental distance ($t_{(2998)} = 6.12$, $p < .001$), and emotional impairment ($t_{(2998)} = 7.02$, $p < .001$). An exception is observed for distress that are more common in Flanders ($t_{(2998)} = -3.34$, $p < .001$), whereas no difference is found with regard to cognitive impairment ($t_{(2998)} = .86$, n.s.). However, the absolute differences between the two countries are rather small, ranging from .10 to .21 for the different scales of the BAT. Such small differences are – although statistically significant – practically speaking not very relevant.

The depression scale of the 4DKL was not administered in the Dutch sample. Yet, validated clinical norms of the are available whereby – after recoding – a value of 2 is considered "moderate" and a value of 5 is considered "strong" (Terluin, 1996, 1998).

Table 64: Norms for the core and secondary symptoms of the BAT for the Netherlands (N = 1,500)

	Total	EX	MD	EI	CI	Sec
Low	≤ 1.55	≤ 1.75	≤ 1.20	≤ 1.20	≤ 1.40	≤ 1.45
Average	1.56 – 2.79	1.76 – 2.99	1.21 – 2.99	1.21 – 2.79	1.41 – 2.60	1.46 – 2.79
High	2.80 – 3.64	3.00 - 3.99	3.00 – 3.99	2.80 – 3.99	2.61 – 3.79	2.80 – 3.59
Very high	≥ 3.65	≥ 4.00	≥ 4.00	≥ 4.00	≥ 3.80	≥ 3.60
M	2.21	2.39	2.18	2.03	2.15	2.11
SD	.82	.88	.96	.94	.88	.84
SE	.02	.02	.02	.02	.02	.02
Range	0.00 – 5.00	0.00 – 5.00	0.00 – 5.00	0.00 – 5.00	0.00 – 5.00	0.00 – 5.00

Note: Total = BAT total-score, EX = BAT exhaustion, MD = BAT mental distance, EI = BAT emotional impairment, CI = BAT cognitive impairment, Sec = BAT secondary distress symptoms, M = mean, SD = standard deviation, SE = standard error of the mean.

A comparison with the average BAT scores of the norm group from Table 59 (p. 106) shows that – without exception – these scores fall into the "high" category. This means that those who score "high" according to Table 64 are comparable to the group with severe burnout .

Table 65 provides an overview of these observed percentages for the "high" and 'very high" scoring categories. It appears that the expected percentage of 5% for "very high" scores is also approached in the Netherlands. As in Flanders, this is somewhat less the case for the "high" percentage, in particular for exhaustion. This means that someone who scores "high" on exhaustion does not belong to the highest 25% but the highest 30% of the Dutch norm group.

Table 65: Observed percentage "high" and "very high" on the BAT for the Netherlands (N = 1,500)

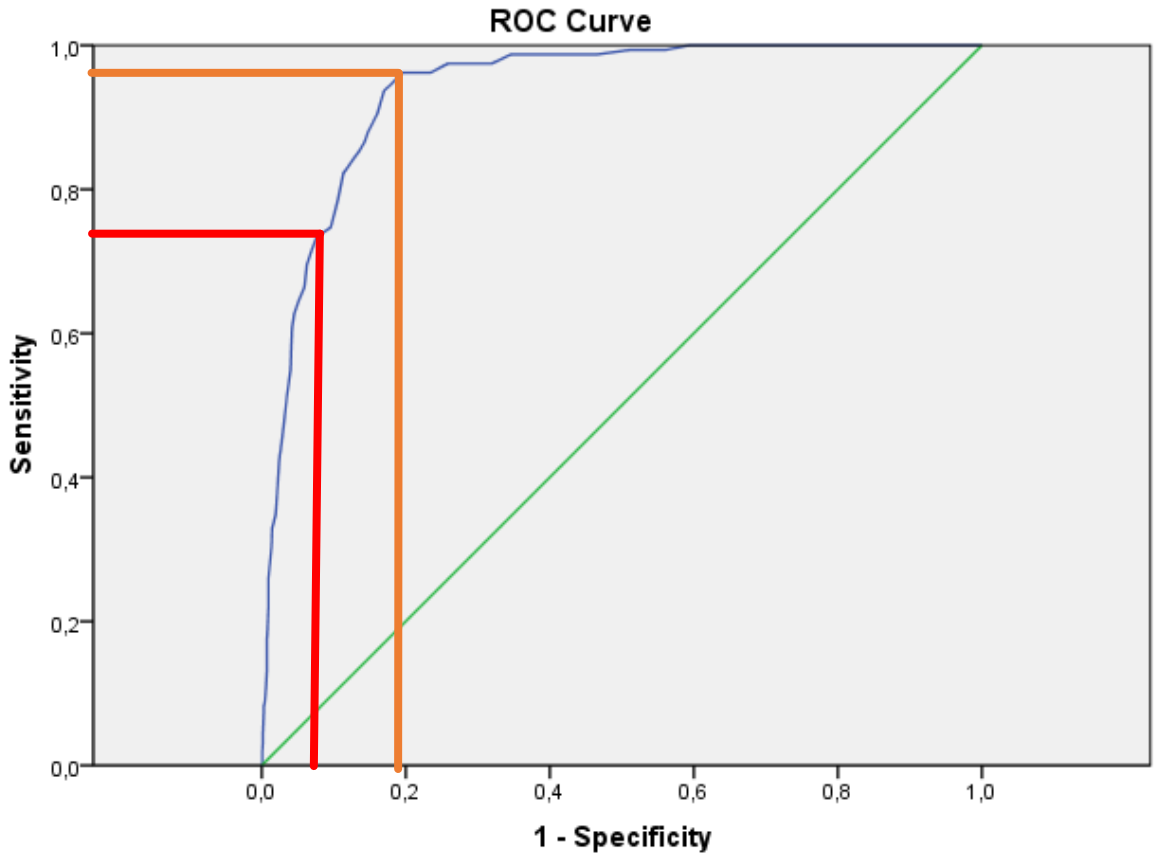
	Total	EX	MD	EI	CI	Sec
% "High"	20.5	25.1	22.4	20.9	21.6	21.3
% "Very high"	5.2	5.6	5.4	5.6	5.6	4.5
% Total	25.7	30.7	27.8	26.5	27.2	25.8

Note: Total = BAT total-score, EX = BAT exhaustion, MD = BAT mental distance EI = BAT emotional impairment, CI = BAT cognitive impairment, Sec = BAT secondary distress symptoms.

3.2. Cut-off values for employees[29]

Figure 9 shows the ROC curve of the total BAT core for Flanders. The cut-off value for orange corresponds to a sensitivity of .94 and a 1-specificity of .17, and that for red to a sensitivity of .73 and a 1-specificity of .08.

Figure 9: ROC curve for the total BAT score in Flanders



In other words, if an orange cut-off value is used, the probability of misclassifying a non-problematic burnout score (the individual is "healthy" or "negative") as problematic (the BAT indicates that the individual is "sick" or "positive") is 17% is. This can be reduced to 8%, but only at the expense of decreasing sensitivity. Sensitivity and specificity are two sides of the same coin: if the specificity (or 1 – specificity) decreases the sensitivity (or 1 – sensitivity) increases, and vice versa. This means that the chance

[29] Please keep in mind that these cut-off values are made for Flanders – they should not be generalized to other groups without caution.

of classifying a problematic burnout score (the individual is "sick" or "positive") as problematic (the BAT correctly indicates that the individual is "sick" or "positive") drops from 94% to 73%. Or reversely, the chance of false positives – that is, those who have problematic score (the individual is "sick" or "positive") but are wrongly classified as non-problematic (the BAT indicates that the individual is "healthy" or "negative") – increases from 6% to 27%.

When using orange cut-offs the meshes of the net are, as it were, larger than when using a more strict red cut-off value. Consequently, the orange group consists of less severe burnout cases than the red group. Put differently, those with an orange score are at-risk for burnout, while those with a red score are most likely to suffer from severe burnout.

Table 66 shows the AUC, the cut-off scores and the corresponding sensitivity and specificity of the BAT scales. Except for mental distance and emotional impairment the AUC is in all other cases greater than or equal to .90. That means that when two random persons – one burnout and one non-burnout case – fill out the BAT, they are classified correctly in 90% of the cases. In other words, the discriminant power of the BAT is excellent.

Table 66: Cut-off scores, sensitivity, specificity and AUC for Flanders

BAT-23	Orange			Red			AUC
	Cut-off	Se	Sp	Cut-off	Se	Sp	
Exhaustion	≥ 3.06	.89	.88	≥ 3.31	.80	.92	.94
Mental distance	≥ 2.50	.75	.77	≥ 3.10	.51	.90	.83
Emotional impairment	≥ 2.10	.89	.72	≥ 2.90	.61	.92	.90
Cognitive impairment	≥ 2.70	.83	.82	≥ 3.10	.68	.94	.93
Total core symptoms	≥ 2.59	.94	.83	≥ 3.02	.73	.92	.94
Secondary symptoms	≥ 2.85	.89	.79	≥ 3.35	.55	.92	.90

Note: Se = sensitivity, Sp = specificity, AUC = area under the curve.

Table 66 shows that the specificity and sensitivity of the orange and red scores is sufficient. Indeed, compared to the orange score, the sensitivity of the red score is somewhat less, which means that the likelihood of "true positives" is somewhat reduced. This applies particularly to mental distance. If a person scores below the orange cut-off score, he or she is in the "green", safe zone. If a person scores between

the orange and red cutoffs, he or she is in the "orange", at-risk zone. And finally, if a person scores above the red cut-off, he or she is in the "red", danger zone.

If the cut-off values from Table 66 are applied to the representative sample of the Flemish workforce, (N = 1,403), 83.1% are in the safe zone (green score), 9.3% are at-risk of burnout (orange score), and 7.6% most likely suffer from severe burnout (red score). Regarding the secondary symptoms; 78.2% is in the safe zone (green score), 13.2% is at-risk of burnout (orange score), 8% is most likely burned-out (red score).

Table 67 provides an overview of scores on various indicators of well-being for the green, orange and red group of Flemish employees from the representative sample (N = 1,403)[30]. The green group scores without exception the most favorable, whereas the red group scores the most unfavorable and the orange risk group]scores in between. There is a significant linear trend; scores on the well-being indicators increase (or decrease) progressively from the green, via the orange to the red group.

Table 67: Well-being of the green, orange and red group; total score core symptoms (Flanders, N = 1,403)

Scale	Mean (M)			$F_{(2, 1400)}$	Eta^2
	Green (N = 1165)	Orange (N = 131)	Red (N = 107)		
BAT-secundary symptoms	2.07	2.87	3.23	272.57***	.48b
Neuroticism	2.76	3.26	3.47	99.68***	.15a
UBOS-exhaustion	2.45	4.36	5.54	537.24***	.54b
UBOS-distance	2.41	4.24	5.09	383.04**	.42b
UBOS-efficacy	5.30	4.29	4.05	136.79***	.18a
OLBI-exhaustion	2.66	3.17	3.53	327.54***	.47b
OLBI-disengagement	2.61	3.22	3.59	321.21***	.43b
DUWAS-working excessively	2.94	3.12	3.20	9.06***	.03
DUWAS-working compulsively	2.52	2.96	3.00	40.66**	.08a
4DKL-depression	1.35	2.19	2.61	286.22***	.32a
4DKL-distress	1.89	2.86	3.25	335.32***	.44b

Note: a Moderate effect; b Strong effect.

With the exception of working excessively (DUWAS), the effect sizes or proportions of explained variance (eta^2) of the various indicators of well-being is moderate or strong.

[30] Those who have been treated for burnout in the past 5 years have been removed from the sample. See also footnote 28.

There is a particularly strong association between the three burnout groups (green, orange, and red) on the one hand and secondary complaints (BAT), exhaustion and mental distance (UBOS, OLBI) and distress (4DKL) on the other hand.

All in all, it can be concluded that the red group is the most problematic one. Those who belong to this group not only score highest on burnout, as measured by two other burnout questionnaires, but also on distress symptoms, neuroticism, working compulsively, and depression. The orange group scores somewhat lower on all of these aspects of well-being, but still clearly higher than the green, healthy group. The opposite was observed for work engagement; the red group exhibits the lowest scores and the green group the highest, with the orange group in the middle.

3.3. Cut-off values for those who do not work

As mentioned earlier, the general version of the BAT was completed by 130 of the 158 Flemish employees with severe burnout (19.2% men and 80.8% women, M_{age} = 43.71, SD = 10.07) and by 756 healthy Flemish employees (70.5 % men and 29.5% women, M_{age} = 39.38, SD = 11.03).

Table 68: Cut-off scores, sensitivity, specificity and AUC for the general version of the BAT in Flanders

BAT-23 general version	Orange			Red			AUC
	Cut-off	Se	Sp	Cut-off	Se	Sp	
Exhaustion	≥ 3.31	.85	.79	≥ 3.81	.55	.91	.88
Mental distance	≥ 2.50	.79	.77	≥ 3.10	.55	.90	.85
Emotional impairment	≥ 3.10	.67	.81	≥ 3.50	.42	.91	.79
Cognitive impairment	≥ 2.70	.92	.73	≥ 3.50	.59	.91	.88
Total core symptoms	≥ 2.59	.97	.83	≥ 3.02	.78	.92	.95
Secondary symptoms	≥ 2.77	.89	.77	≥ 3.23	.59	.91	.90

Note:, Se = sensitivity, Sp = specificity, AUC = area under the curve.

The discriminant power of the general version of the BAT is excellent because the AUC is well above .60 for all scales. Table 68 also shows that the specificity and sensitivity of the orange and red cut-off scores is sufficient. Nevertheless, the sensitivity of the red score is somewhat lower, which means that the likelihood of "true positives" is slightly reduced. This applies particularly to the subscales of the BAT and not to the total score. Therefore, the cut-off scores of the subscales of the general version of the BAT should be used with caution.

4. Norms and cut-off values for the BAT-12

As for the BAT-23, statistical norms and clinically validated cut-off values have been calculated for the BAT-12. The same procedure was followed as outlined above for the BAT-23, using a representative, healthy sample of the Flemish and Dutch working population, and a group with severe burnout complaints, respectively. For more details about this group with severe burnout complaints, see pp. 105-106.

Based on statistical norms, it is possible to classify BAT-12 scores into four categories: (1) "low"; (2) "average"; (3) "high" and (4) "very high". For more information about the statistical norms, see section 2.1 (p. 103).

Based on clinical cut-off scores, individuals can be grouped into three categories: (1) without burnout (green); (2) at risk for burnout (orange) and (3) likely to be burned-out (red). For more information on the clinical cutoffs, see sections 2.2 (pp. 104-107) and 3.2 (pp. 112-115).

4.1. Statistical norms

Tables 69 and 70 provide an overview of the statistical norms for Flemish and Dutch workers, respectively, based on a representative sample of the working population.

Table 69: Statistical norms for the BAT-12 for Flanders (N = 1,500)

BAT-12	Total	EX	MD	EI	CI
Low	≤. 1.50	≤. 1.66	≤. 1.00	≤. 1.00	≤. 1.66
Average	1.51 – 2.35	1.67 – 2.99	1.01 – 2.65	1.01 – 2.00	1.67 – 2.33
High	2.36– 3.17	3.00 – 3.99	2.66 – 3.99	2.01 – 3.00	2.34 –3.32
Very high	≥ 3.18	≥ 4.00	≥ 4.00	≥ 3.01	≥ 3.33
M	2.02	2.26	1.98	1.73	2.12
SD	.66	.86	.90	.74	.70
SE	.02	.02	.02	.02	.02
Range	1.00-5.00	1.00-5.00	1.00-5.00	1.00-5.00	1.00-5.00

Note: Total = total score (12 items), EX = BAT-12 exhaustion, MD = BAT-12 mental distance, EI = BAT-12 emotional impairment, CI = BAT -12 cognitive impairment, M = mean, SD = standard deviation; SE = standard measurement error.

Due to the specific distribution of scores in the normative sample, the effective percentages of respondents may deviate from the theoretically expected percentages. Theoretically speaking, 5% should fall into the "very high" category and 20% into the "high" category. Particularly for the short subscales of the BAT-12 which

only consist of 3 items, deviations from the expected percentages may be relatively large compared to the longer subscales of the BAT-23. Tables 70 and 72 indeed show that the group of respondents who score "high" or "very high" on the BAT-12 scores reasonably corresponds to the expected size of 20% resp. 5%.

Table 70: Effective percentage 'high' and 'very high' on the BAT-12 in Flanders (N = 1,500)

BAT-12	Total	EX	MD	EI	CI
% 'High'	19.3	18.5	21.3	19.2	33.3
% 'Very high'	5.7	5.8	4.2	4.3	6.3
% Total	25.0	24.3	25.5	23.5	39.6

Note: Total = total score (12 items), EX = BAT-12 exhaustion, MD = BAT-12 mental distance, EI = BAT-12 emotional impairment, CI = BAT -12 cognitive impairment.

Table 71: Statistical norms for the BAT-12 for the Netherlands (N = 1,500)

BAT-12	Total	EX	MD	EI	CI
Low	≤ 1.50	≤ 1.66	≤ 1.00	≤ 1.00	≤ 1.33
Average	1.51 – 2.79	1.67 – 2.67	1.01 – 2.99	1.01 – 2.67	1.34 – 2.67
High	2.80 – 3.66	2.68 - 3.99	3.00 – 3.99	2.68 – 3.59	2.68 – 3.66
Very high	≥ 3.67	≥ 4.00	≥ 4.00	≥ 3.60	≥ 3.67
M	2.16	2.35	2.14	1.98	2.17
SD	.84	.93	.99	.95	.89
SE	.02	.02	.03	.02	.02
Range	1.00 – 5.00	1.00 – 5.00	1.00 – 5.00	1.00 – 5.00	1.00 – 5.00

Note: Total = total score (12 items), EX = BAT-12 exhaustion, MD = BAT-12 mental distance, EI = BAT-12 emotional impairment, CI = BAT -12 cognitive impairment, M = mean, SD = standard deviation; SE = standard measurement error.

Table 72: Effective percentage 'high' and 'very high' on the BAT-12 in the Netherlands (N = 1,500)

BAT-12	Total	EX	MD	EI	CI
% 'High'	20.1	25.5	21.4	17.0	17.3
% 'Very high'	5.0	7.6	6.9	7.1	7.9
% Total	25.1	33.1	28.3	24.1	25.4

Note: Total = total score (12 items), EX = BAT-12 exhaustion, MD = BAT-12 mental distance, EI = BAT-12 emotional impairment, CI = BAT -12 cognitive impairment.

However, deviating percentages are found for cognitive impairment in Flanders (Table 70) and exhaustion in the Netherlands (Table 72). This means, for example, that someone who scores "high" on cognitive impairment in Flanders does not belong to the highest 25% but to the highest 40% of the norm group. And someone who scores "high" on exhaustion in the Netherlands actually belongs to the highest scoring 33% of

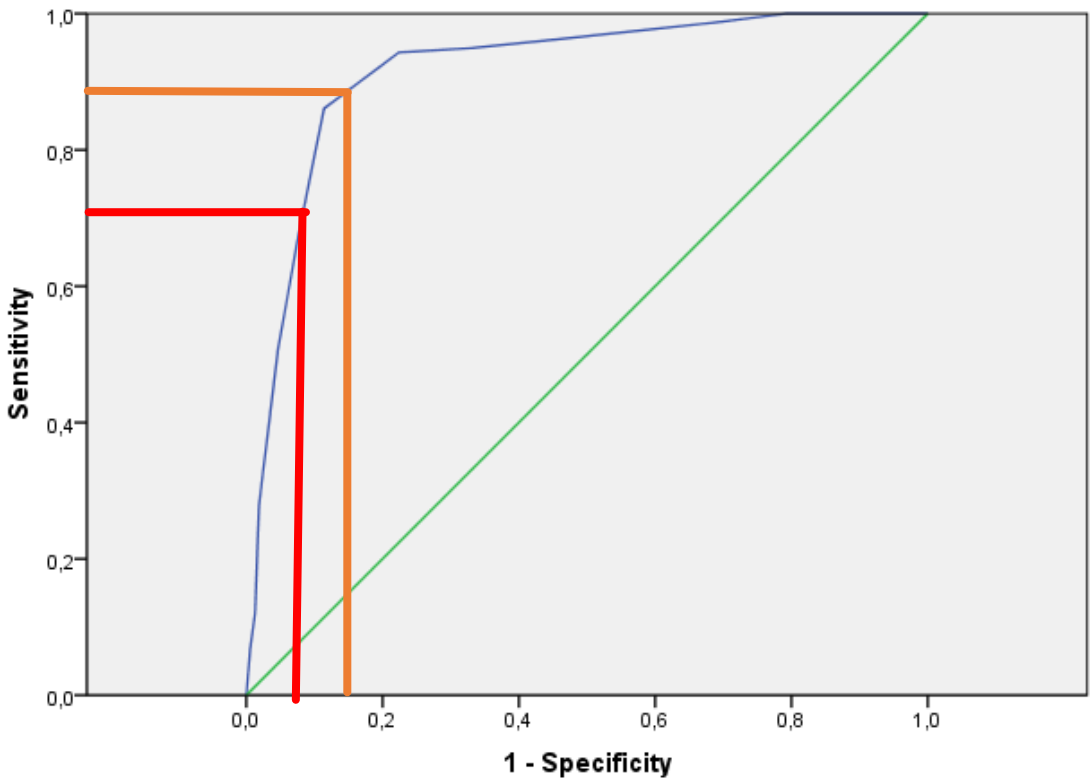
the norm group. In other words, the classification of subscales gives a worse picture than that of the total score.

4.2. Cut-off values

The procedure followed to determine the cut-off values is set out in section 2.2 (pp. 104-107), whereby we used a group of 158 Flemish employees with serious burn-out complaints (see Table 59, p. 106).

Figure 10 shows the ROC curve of the BAT-12 for Flanders. The cut-off value for the orange risk group has been determined with a sensitivity of .90 and a 1-specificity of .17, and that for the red burnout group with a sensitivity of .72 and a 1-specificity of .05.

Figure 10: ROC-curve fort the total score of the BAT-12 for Flanders



In other words, if one uses an orange cut-ff value, the probability of wrongly classifying a non-problematic burnout score (i.e., someone is "healthy" or "negative") as

problematic is 17%. In that case the BAT incorrectly indicates that the person is "sick" or "positive". This can be reduced to just 5%, but at the expense of sensitivity. After all, when the specificity (or 1 - specificity) decreases, the sensitivity (or 1 - sensitivity) increases correspondingly, and vice versa. This means that the likelihood that a problematic burnout score (i.e., someone is "sick" or "positive") is indeed correctly classified as problematic (and hence the BAT indicates "sick" or "positive") decreases from 90% to 72%. Or conversely, the likelihood of false positives – that is, those who have a problematic score (and are "sick" or "positive") and classified wrongly as non-problematic (the BAT indicates "healthy" or "negative") – increases from 5% to 28%.

Table 73 shows the Area Under the curve (AUC), the cut-off values as well as the associated sensitivity and specificity of the BAT-12. The AUC is in all cases greater than or equal to .84. This means that when two random individuals – one with one without serious burnout complaints – fill in the BAT, they are correctly classified in at least 84% of the cases. For the total score that is even 93%.

Table 73: Cut-off values, sensitivity, specificity, and AUC for Flanders

BAT-12	Orange			Red			AUC
	Cut-off	Se	Sp	Cut-off	Se	Sp	
Exhaustion	≥ 3.17	.86	.89	≥ 3.50	.68	.92	.92
Mental distance	≥ 2.17	.85	.67	≥ 3.17	.54	.91	.84
Emotional impairment	≥ 2.17	.76	.80	≥ 2.83	.56	.92	.85
Cognitive impairment	≥ 2.83	.78	.86	≥ 3.17	.58	.95	.89
Total BAT-12	≥ 2.54	.90	.83	≥ 2.96	.72	.91	.93

Note: Se = sensitivity, Sp = specificity; AUC=Area under the curve.

In other words, the discriminating power of the BAT-12 is excellent, also in comparison with the BAT-23 (Table 66, p. 113). The discriminatory power of the subscales of the short BAT is only a fraction below that of the BAT-23 and differs by 4% at most (for cognitive impairment). For the total score the difference is negligible and amounts to only 1% (.95 for the BAT-23 vs. .93 for the BAT-12). This means that the BAT-12 can be used just as well as the longer version to screen employees for the occurrence of serious burn-out complaints.

If the cut-off values for the BAT-12 from Table 74 are applied to the representative sample of Flemish employees, (N = 1,403)[31], 7.3% has a risk of burnout (orange score) and 9.3% suffers from serious burnt out complaints (red score), while 83.4% are in the safe zone (green score). Based on the cut-off values of the BAT-23, a comparable group (83.1%) is classified in the safe zone (green score), but the prevalence of the group at-risk (orange score) is estimated 2% less and that of the burned-out group (red score) is estimated 1.7% higher. In other words, compared to the BAT-23, the BAT-12 provides a somewhat higher estimate for the problematic (red) group and a somewhat lower estimate for the (orange) risk group. However, if both groups are combined, the prevalence estimates of the two BAT versions hardly differ (16.9% for the BAT-23 vs. 16.6% for the BAT-12).

Table 74: Well-being of the green, orange and red groups (Flanders, N = 1,403)

Scale	Mean (M)			$F_{(2, 1400)}$	Eta^2
	Green (N = 1160)	Orange (N = 107)	Red (N = 136)		
BAT-secondary symptoms	2.08	2.85	3.20	273.68***	.41[b]
Neuroticism	2.76	3.24	3.41	93.83***	.12[a]
UWES	3.78	3.22	2.64	172.52***	.20[a]
UBOS-exhaustion	2.45	4.38	5.23	498.21***	.42[b]
UBOS-distance	2.40	4.28	4.92	393.13**	.36[b]
UBOS-competence	5.30	4.43	4.03	134.43***	.16[a]
OLBI-exhaustion	2.66	3.17	3.43	302.55***	.30[b]
OLBI-disengagement	2.61	3.22	3.74	318.26***	.31[b]
DUWAS-excessive	2.93	3.16	3.20	12.31***	.02
DUWAS-compulsive	2.51	2.93	3.05	46.76***	.06[a]
4DKL-depression	1.36	1.97	2.64	296.20***	.30[b]
4DKL-distress	1.89	2.77	3.19	327.66***	.32[b]

Note: [a] Moderate effect; [b] strong effect.

Table 74 provides an overview of mean well-being scores of the green, orange and red groups of Flemish employees from the representative sample (N = 1,403). Without exception, the green, healthy group scores the most favorably and the red, burned-out group the most unfavorably, while the orange risk group is always found in between. For all well-being indicators there is a significant linear trend from the green, via the orange to the red group. With the exception of excessive working (DUWAS),

[31] Those who have been treated for burnout complaints in the past 5 years were removed. See also footnote 28.

the effect sizes, or the proportion of explained variance (eta^2), of the indicators of well-being are moderate or strong.

There is a strong relationship between the three burnout groups on the one hand and secondary complaints (BAT), exhaustion and distance/disengagement (UBOS, OLBI), and distress (4DKL) on the other. This pattern fully matches that of the full BAT-23 (see Table 67, p. 114). This means that the well-being profile of the green, orange and red group does not differ, irrespectively when based on the BAT-23 or the BAT-12.

4.3. Cut-off values for those who do not work

For determining cut-off values of the context-free, general version of the BAT-12, a group of working people was used for Flanders (N = 756; 70.5% men and 29.5% women, M_{age} = 39.38, SD = 11.03) supplemented by a part of the group with severe burnout complaints (N = 130: 19.2% men and 80.8% women, M_{age} = 43.71, SD = 10.07). Unfortunately, the mental distance items were not administered in the healthy group, so that no cut-off values for the total, general BAT-12 are available. For more information about both group (see p. 108 and p. 115).

The discriminating power of the subscales of the general version of the BAT is good because the AUC is well above .60 in all cases. Furthermore, Table 74 shows that the specificity and sensitivity of the orange scores is sufficient, as is the specificity of the red scores. However, the sensitivity of the red scores is below par, which means that the likelihood of "true positives" is greatly reduced. It is therefore recommended that only the cut-off values for the orange scores of the general BAT be used.

Table 74. Cut-off values, sensitivity, specificity and AUC for the general version of the BAT-12 in Flanders

BAT-12 general version	Orange			Red			AUC
	Cut-off	Se	Sp	Cut-off	Se	Sp	
Exhaustion	≥ 3.50	.80	.79	≥ 4.17	.28	.94	.85
*Mental distance**	≥ 2.17	.85	.67	≥ 3.17	.54	.91	.84
Emotional impairment	≥ 2.83	.75	.69	≥ 3.83	.22	.95	.77
Cognitive impairment	≥ 2.83	.78	.73	≥ 3.50	.48	.91	.87

Note: Se = sensitivity, Sp = specificity; AUC=Area under the curve; * identical with the work related version (see Table 71).

5. About the use of norms and cut-off values

The BAT can be used for individual assessment of employees as well as for screening (i.e. identifying those who are at-risk or suffering from burnout) and benchmarking (i.e. comparing burnout levels with the national workforce). In addition, the general version of the BAT can be used for monitoring those who have not been working for a longer period of time (a few months). Below, a more detailed description is given of how norms and cut-off values can be used in each of these cases.

5.1. Full or short version?

The psychometric properties of the shortened version of the BAT, which consists of 12 questions, are similar to that of the full BAT, which consists of 23 questions (see Part III). This means that, technically speaking, the shortened version can be used just as well as the full version. This is especially true when the BAT is used for benchmarking, whereby the level of burnout complaints of a group of employees is compared to that of the average Flemish or Dutch employee.

However, it is better to use the full version for individual assessment and for screening purposes with the aim of identifying those wo are at risk for burnout. First, because the scores on the full version are more accurate (i.e., more differentiated), simply because the full version contains more questions, resulting in a finer-grained score distribution. Secondly, the discriminative power of the shortened version is slightly less than that of the full version. This means that burnout cases and non-burnout cases are less adequately distinguished from each other. Both reasons especially apply for the subscales of the short version; the total score can be used for individual assessment and screening if one is not interested in a breakdown into different core symptoms.

5.2. Individual assessment

The so-called traffic light model can be used for individual assessment. That is, the cut-off values each correspond with one of the colors of a traffic light (green orange, red). These cut-off values are displayed in Tables 66 (BAR-23) and 73 (BAR-13) are displayed

For the total score on the BAT in Flanders this means the following:

- Green: the score is lower than 2.58 (2.53) – no burnout exists;
- Orange: the score is higher than or equal to 2.59 (2.54) and less than 3.02 (2.96) – there is a risk of burning out;
- Red: the score is higher or equal to 3.02 (2.96) – burnout is most likely.

Note: Cut-off values of the BAT-12 between parentheses.

In the case of an orange or red score on the total BAT, scores on the four subscales can subsequently be inspected for a more detailed picture. For example, it may turn out that someone who is at-risk for burnout (orange) scores red for emotional impairment, but green for mental distance, or the other way around.

Secondary symptoms (BAT) and depression (4DKL) can be considered in a similar way. For the interpretation of the depression score of the 4DKL, see Terluin (1998). As indicated above, it is not recommended to use sub-scale scores of the BAT-12f or individual diagnostic use.

Because clinically validated cut-offs are not yet available for the Netherlands, the statistical norms from Tables 64 (BAT-23) and 71 (BAT-12) can be used for the time being. Based on a comparison of the Flemish cut-off values (Table 66 for the BAT-23 and Table 73 for the BAT-12) with the Flemish statistical norms (Tables 60 and 62 for the BAT-23 and Table 69 for the BAT-12), it can be assumed that certainly those who score "very high" – and probably also a large proportion of those who score "high" – run a risk of burnout or may actually suffer from severe burnout.

A confidence interval can be taken into account for interpreting individual scores more accurately. After all, a particular uncertainty margin or confidence interval exists around the observed score. The standard error of measurement (SE) – which for all scales of the BAT equals .02 or .03 (see Tables, 60, 62 and 64 for the BAT-23 and Tables 69 and 71 for the BAT-12) – can be used for estimating this interval. A 95% confidence interval can be calculated as follows. Suppose someone scores 3.00 on the core symptoms of the BAT; the 95% confidence interval is then 3.00 ± 1.96 x .02. That means that the "true" score lies with a probability of 95% between 2.96 and 3.04.

Of course, no burnout diagnosis can be drafted solely on the basis of the BAT, or any other self-report burnout questionnaire for that matter. This requires a more extensive assessment, including an anamnestic and diagnostic interview by a trained professional. However, the BAT can be a valuable tool to use in the assessment process for estimating the individual's level of burnout symptoms.

5.3. Screening

For organizations it is important to know how many employees run a risk of burnout and how many might actually suffer from burnout. For Flanders the cut-off values from Table 66 (BAT-23) and Table 73 (BAT-12) can be used for this purpose; by applying these cut-offs the percentage of employees can be estimated who are in the green (no burnout), orange (at-risk of burnout) and red (most likely burned-out) groups. Departments, teams, wards, jobs, age groups, locations, and so on may be compared on the basis of this classification. A screening may identify individuals or groups to which specific preventive measures may be targeted.

Because no clinically validated cut-off values are yet available for the Netherlands, for the time being the statistical norms from Table 64 can be used for screening purposes (see also under benchmarking).

5.4. Benchmarking

Benchmarking involves comparing burnout levels of employees of a particular organization (or part thereof) with that of the average Flemish or Dutch employee. By using the statistical norms from Tables 60 and 62 (Flanders) and 64 (the Netherlands) for the BAT-23, and Table 69 (Flanders) and Table 71 (the Netherlands) for the BAT-12, the average level of burnout of an organization – or a part thereof – can be classified as "low", "average", "high", or "very high" compared to national workforce. In contrast to screening, the focus is not on identifying (potential) burnout cases, but rather on comparing levels of burnout with a reference group that may serve as a meaningful standard.

Benchmarking can be important because it may guide an organization's HRM and occupational health policies, for instance. Most organizations conduct employee

satisfaction surveys or working conditions surveys on a regular basis, which also may include burnout. These surveys such as, for instance, the Energy Compass, may identify potential drivers (i.e., job demands and lack of job resources) or burnout (Schaufeli, 2017b). Hence, based on the results of these surveys HRM- and occupational health policies may be developed to prevent burnout or specific risk groups may be targeted.

5.5. Monitoring

It is imperative to examine the extent to which burnout decreases during the course of a treatment, counseling or return-to-work program. In many cases, the regular version of the BAT will not be suitable for this purpose because the persons involved may have been absent from work for a longer period of time. For such cases, the general version of the BAT has been developed, for which cut-off values were calculated for Flanders (Table 68 for the full BAT and Table 74 for the BAT-12). For the Netherlands these will follow later.

The traffic light model may also be used for monitoring as well because it is important to know whether or not BAT scores droppe from red to orange, or from orange to green. However, it is also important to note that meaningful changes in burnout levels may occur *within* the green, orange, or red groups. Consider, for instance, the case where an orange score of 3.0 drops to 2.6 on the BAT-23. This may indicate a relevant decrease in level of burnout – despite the fact that 2.6 is still in the orange range – because the 95% confidence intervals of both scores (2.96 - 3.04 or 2.56 - 2.64) do *not* overlap. This mean that there is indeed a "true" difference between both scores. In other words, when monitoring it is also important to look at changes in average scores in addition to using the traffic light.

6. Conclusion

The BAT can be used for individual assessment, as well as for screening and benchmarking in organizations. In addition to the standard BAT, which is intended for those who are at work or who have been working until recently (a few weeks to about one month), there is also a general version that is intended for assessing and monitoring those who have not been working for some time (a few months or so).

In addition to the total score on the BAT, the scores on the four scales (i.e., exhaustion, mental distance, and emotional and cognitive impairment) can be used as well. For individual assessment, the use of subscale-scores of the BAT-12 is *not* recommended due to their low specificity as shown by the ROC-analysis. The BAT also contains a scale for secondary distress symptoms that often occur together with burnout, but are not specific for burnout. Finally, it is advised to use a short depression questionnaire as well, such as that of the 4DKL (Terluin et al., 2006), particularly for individual assessment.

In most cases, the overall score of the BAT-12 will be sufficient for screening organizations. However, for individual assessment a more differentiated picture is usually needed that would also include the four subscales of the BAT-23 as well as secondary symptoms and depressed mood.

It is possible to determine the relative level of burnout on the basis of norms that are based on representative samples of the Flemish or Dutch workforce (benchmarking). Accordingly, each score can be classified as low (below the 25th percentile), average (25th - 75th percentile), high (75 - 95th percentile) or very high (above the 95th percentile) in relation to average Flemish or Dutch employee. Because BAT scores differ slightly between countries, separate statistical norms have been calculated for Flanders and the Netherlands. As a rule, benchmarking will take place in organizations, whereby the entire organization, or parts thereof, are compared with the national workforce. However, scores of individual employees might also be compared with the reference group.

The ROC analyses show that the discriminant power of the BAT is excellent. This means that the BAT is able to make a distinction between those with and without severe burnout symptoms.

Cut-off scores have been calculated that allow to make a distinction between burnout cases and non-cases. The burned-out group that was used to calculate these cut-off values was selected on the basis of a number of inclusion and exclusion criteria that were derived from the Dutch *Multidisciplinary Guideline on Stress and Burnout* and the diagnostic criteria for Exhaustion Disorder by the *Swedish National Board of Health and Welfare*. A burned-out group is currently only available for Flanders, so that no cut-off values for the Netherlands are available yet.

Using the established cut-off values, the total scores on core symptoms and the secondary symptoms of the BAT can be divided into three groups, that are labeled as traffic lights:

- Green: there is no burnout;
- Orange: there is a risk of burnout;
- Red: burnout is very likely.

A red score means that the specificity is greater than or equal to .90, which corresponds to a maximum of 10% misclassifications; that is, a burnout case that is overlooked.

Our findings show that the red group not only scores highest on other burnout questionnaires, but also on psychological distress, neuroticism, working compulsively, and depressed mood. The orange group scores somewhat lower on all aspects of well-being, but still clearly higher than the green, healthy group. The opposite applies for work engagement; the red group scores the lowest and the green group the highest, with the orange group in the middle.

Also for the general version of the BAT orange and red cut-off values for Flanders are available. The discriminant power of the full and short general version is very good too. However, the sensitivity of the red scores of the BAT subscales is somewhat less, so that caution is warranted when interpreting the results.

Literature[32]

Alarcon, G.M. (2011). A meta-analysis of burnout with job demands, resources and attitudes. *Journal of Vocational Behavior, 79*. 549-562.

Alarcon, G., Eschleman, K.J., & Bowling, N. (2009). Relationships between personality variables and burnout: A meta-analysis. *Work & Stress, 23,* 244–263.

Baillien, E., De Witte, H., Notelaers, G., & Neyens, I. (2008). De ontwikkeling van de RATOG: een screeninginstrument voor de preventie van pesterijen op het werk [The development of RATOG: A screening instrumnet for the pervention of bullying at work]. *Gedrag & Organisatie, 21*, 254-277.

Bakker, A.B., & Demerouti, E. (2016). Job Demands – Resources theory: taking stock and looking forward. *Journal of Occupational Health Psychology, 22*, 273–285.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Beck, A.T, Steer, R.A., & Brown, G.K. (1996). *Beck Depression II. Test manual*. San Antonio, TX: The Psychological Corporation.

Boudreau, R.A., Cahoon, S.L., & Wedel, R.J. (2006). The pandemic from within: two surveys of physician burnout in Canada. *Canadian Journal of Community Mental Health, 25*, 71–88.

Boudreau, R.A., Boudreau, W.F., & Mauthe-Kaddoura, A.J. (2015). *From 57 for 57: a bibliography of burnout citations.* Poster presented at the 17th Conference of the European Association of Work and Organizational Psychology (EAWOP): Oslo, Norway.

Brenninkmeijer, V., & Van Yperen, N. (2003). How to conduct research on burnout: advantages and disadvantages of a unidimensional approach to burnout. *Occupational and Environmental Medicine, 60*, 16-21. (Suppl 1).

[32] The publications that have been (co-)authored by Schaufeli can be downloaded from www.wilmarschaufeli.nl.

Bresó, E., Salanova, M., & Schaufeli, W.B (2007). In search of the 'third dimension' of burnout: Efficacy or inefficacy? *Applied Psychology: An International Review, 56,* 460-478.

Burisch, M. (2017). *Hamburg Burnout Inventory (HBI). Background and some early results*. Downloaded November 11th 2018 from https://www.burnout-institut.eu/fileadmin/user_upload/HBI-REP_V2_FINAL.pdf.

Byrne, B, (2010). *Structural equation modeling with AMOS* (3rd. Ed). Mahwah, NJ: Lawrence Erlbaum.

Clark, M.A, Michel, J.S., Zhdanova, L., Pui, S.Y., & Baltes, B.B. (2016). All work and no play? A meta-analytic examination of the correlates and outcomes of workaholism. *Journal of Management, 42,* 1836-1873.

Crawford, E.R., Lepine, J.A., & Rich, B.L. (2010). Linking job demands and resources to employee engagement and burnout: a theoretical extension and meta-analytic test. *The Journal of Applied Psychology, 95*, 834–848.

Christensen, K.B., Kreiner, S., & Mesbah, M. (2013). *Rasch models in health.* London: Wiley.

De Beer, L.T., & Bianchi, R. (2019). Confirmatory Factor Analysis of the Maslach Burnout Inventory. European *Journal of Psychological Assessment*, 35, 217-224.

De Bie, D. (2018). Zien anderen mij opbranden? Een peer-evaluatie studie van de BAT-vragenlijst. Masterproef aangeboden tot het verkrijgen van de graad van Master of Science in de Psychologie, KU Leuven.

Deligkaris, P., Panagopoulos, E., Montgomery, A., & Masoura, E. (2014). Job burnout and cognitive functioning: A systematic review. *Work & Stress, 28,* 107-123.

Demerouti, E., Bakker, A.B., Vardakou, I., & Kantas, A. (2003). The convergent validity of two burnout instruments: A multitrait-multimethod analysis. *European Journal of Psychological Assessment, 18*, 296-307.

De La Fuente, E.I., Lozano, L.M., García-Cueto, E., Luis, C. S., Vargas, C., Cañadas, G.R., & Hambleton, R.K. (2013). Development and validation of the Granada Burnout Questionnaire in Spanish police. International, 216–225.

De Vente, W., Olff, M., Van Amsterdam, J.C.G., Kampuis, J.H., & Emmelkamp P.M.G. (2013). Physiological differences between burnout patients and healthy controls: blood pressure, heart rate, and cortisol responses. *Occupational and Environmental Medicine, 60*, i54-i61.

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing, 62*, 107–115.

Federale Overheidsdienst (FOD) (2017). *Tool voor de vroegtijdige opsporing van burn-out [Tool for the early detection of burnout]*. Downloaded November 11th, 2018; http://www.werk.belgie.be/publicationDefault.aspx?id=44193.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th Ed.) Gage: London.

Fischer, J.E., Bachmann, L.M., & Jaeschke, R. (2003). A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. *Intensive Care Medicine, 29*, 1043-1051.

Formel, C., & Larcker, D.F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*, 39-50.

Gil-Monte, P., & Faúndez, V. (2011). Psychometric properties of the 'Spanish Burnout Inventory' in Chilean professionals working with physical disabled People. *The Spanish Journal of Psychology, 14*, 441-451.

González-Romá, V., Schaufeli, W.B., Bakker, A., & Lloret, S. (2006). Burnout and engagement: independent factors or opposite poles? *Journal of Vocational Behavior, 68*, 165-174.

Glise, K., Hadzibajramovic, E., Jonsdottir, I. H., & Ahlborg, G. (2010). Self-reported exhaustion: a possible indicator of reduced work ability and increased risk of

sickness absence among human service workers. *International Archives of Occupational and Environmental Health, 83*, 511–520.

Habibazadeh, F., Habibazadeh, P., Yadollahie, M. (2016). On determining the most appropriate test cut-off value: the case of tests with continuous results. *Biochemical Medicine, 26*, 97–307.

Hagenaars J.A., & McCutcheon, A.L. (2009). *Applied latent class analysis*. Cambridge University Press.

Hartley, J. (2013). Some thoughts on Likert-Scales. *International Journal of Clinical and Health Psychology, 13*, 83-86.

Hagquist, C., Bruce, M., & Gustavsson, J.P. (2009). Using the Rasch model in nursing research: An introduction and illustrative example. *International Journal of Nursing Studies, 46,* 380-93.

Hu, L., & Bentler, P. (1995). Evaluating Model Fit. In R.H. Hoyle (Ed.), *Structural Equation Modeling. Concepts, Issues, and Applications* (pp. 76–99). Thousand Oaks, CA: Sage.

Jonsdottir, I.H., Nordlund, A., Ellbin, S., Ljung, T., Glise, K., Währborg, P., Sjörs, A., & Wallin, A. (2017). Working memory and attention are still impaired after three years in patients with stress-related exhaustion. *Scandinavian Journal of Psychology, 58*, 504-509.

Kleijweg, J.H.M., Verbraak, M.J.P.M., & Van Dijk, M.K. (2012). The clinical utility of the Maslach Burnout Inventory in a clinical population. *Psychological Assessment, 25,* 435-411.

Kristensen, T.S., Borritz, M., Villadsen, E., & Christensen, K.B. (2005). The Copenhagen Burnout Inventory: a new tool for the assessment of burnout. *Work & Stress, 19*, 192–207.

Lastovkova, A., Carder, M., Rasmussen, H.M., Sjoberg, L., et al. (2017). Burnout syndrome as an occupational disease in the European Union: an exploratory study. *Industrial Health, 56,* 160–165.

Maslach, C. (1986). Stress, burnout and workaholism. In R.R. Killberg, P.E. Nathan, & R.W. Thoreson (Eds.), *Professionals in distress: Issues, syndromes and solutions in psychology* (pp. 53-73). Washington (DC): American Psychological Association.

Maslach, C., & Jackson, S.E (1981a). The measurement of experienced burnout. *Journal of Occupational Behavior, 2*, 99–113.

Maslach, C., & Jackson, S.E. (1981b). *Maslach Burnout Inventory Manual*. Palo Alto, CA: Consulting Psychologists Press.

Maslach, C., & Jackson, S.E. (1986). *Maslach Burnout Inventory Manual*. (2nd. Ed.) Palo Alto, CA: Consulting Psychologists Press.

Maslach, C., Jackson, S.E., & Leiter, M.P. (1996). *Maslach Burnout Inventory Manual (3rd. Ed.)*. Palo Alto, CA: Consulting Psychologists Press.

Maslach, C., Jackson, S.E., Leiter, M.P., Schaufeli, W.B. & Schwab, R.L (2017). *Maslach Burnout Inventory Manual (4rd. Ed.)*. Palo Alto, CA: Mind Garden.

Maslach, C., Schaufeli, W.B., & Leiter, M.P. (2001). Job Burnout. *Annual Review of Psychology, 52*, 397-422.

Meijman, Th., & Schaufeli, W.B. (1996). Psychische vermoeidheid en arbeid [Menal fatige and work]. *De Psycholoog, 31*, 236-241.

Moreno-Jimenez, B., Barbaranelli, C., Galvez-Herrer, M., & Garrosa, E. (2012). The physician burnout questionnaire: a new definition and measure. *Testing, Psychometrics, Methodology in Applied Psychology, 19*, 325–344.

Nielsen T., & Kreiner S. (2011). Reducing the Item number to obtain same-length self-assessment scales: A systematic approach using result of graphical loglinear Rasch modeling. *Journal of Applied Measurement,12*, 310-23.

Nielsen T., & Kreiner S. (2013). Improving items that do not fit the Rasch model. In: K.B. Christensen, S. Kreiner, & M. Mesbah (Eds*). Rasch models in health* (pp. 317-34). London: Wiley.

Notelaers, G., De Witte, H., & Van Veldhoven, M. (2005). Kengetallen of grenswaarden voor psychische vermoeidheid in stressonderzoek. Een vergelijkende studie met de data van de 'Vlaamse Werkbaarheidsmonitor' [Key figures of cut-of values for mental fatigue in stress research. A comparative study using data of the 'Flemish Workability Monitor']. *OverWerk, 4*, 187-193.

Nunnally, J.C., & Bernstein, I.H. (1994). *Psychometric theory* (3rd. Ed.). New York: McGrawHill.

Pines, A., Aronson, E., & Kafry, D. (1981). *Burnout: from tedium to personal growth*. New York: The Free Press.

Reijseger, G., Schaufeli, W.B., Peeters, M.C.W., Taris, T.W. van Beek I., & Ouweneel, E. (2013). Watching the paint dry: validation of the Dutch Boredom Scale. *Anxiety, Stress & Coping, 26*, 508-525.

Salmela-Aro, K., Rantanen, J., Hyvönen, K., Tilleman, K., & Feldt, T. (2011). Bergen Burnout Inventory: reliability and validity among Finnish and Estonian managers. *International Archives of Occupational and Environmental Health, 84*, 635–45.

Schaffner, A.K. (2016). *Exhaustion. A history*. New York: Columbia University Press.

Schaufeli, W.B. (2015). Van burnout naar bevlogenheid. Werk en welbevinden in Nederland [From burnout to engagement. Work and well-being in the Netherlands]. *M&O, 69*, 15-31.

Schaufeli, W.B. (2017a). Burnout: a short socio-cultural history. In S. Neckel, A.K. Schaffner & G. Wagner (Eds.), *Burnout, fatigue, exhaustion: an interdisciplinary perspective on an modern affliction* (pp. 105-127). Cham: Palgrave MacMillan.

Schaufeli, W.B. (2017b). The Job Demands-Resources model; A 'how to' guide to increase work engagement and prevent burnout. *Organizational Dynamics, 46*, 120-132.

Schaufeli, W.B. (2018). Burnout: feiten en fictie [Burnout: Facts and fiction]. *De Psycholoog, 53*, 9, 10-20.

Schaufeli, W.B. & Bakker, A.B. (2004a). Bevlogenheid: een begrip gemeten [Work engagement: The measurement of a construct]. *Gedrag & Organisatie, 17,* 89-112.

Schaufeli, W.B., & Bakker, A.B. (2004b). Job demands, job resources and their relationship with burnout and engagement: A multi-sample study. *Journal of Organizational Behavior, 25*, 293-315.

Schaufeli, W.B. Bakker, A.B. , Schaap, C. Kladler, A., & Hoogduin, C.A.L. (2001) On the clinical validity of the Maslach Burnout Inventory and the Burnout Measure. *Psychology & Health, 16*, 565-582.

Schaufeli. W.B., & De Witte, H. (2017). Work engagement in contrast to burnout: real and redundant! *Burnout Research, 5*, 58-60.

Schaufeli, W.B., & Enzmann, D. (1998). *The burnout companion to study and research: a critical analysis*. London: Taylor & Francis.

Schaufeli, W.B., Leiter, M.P., & Maslach, C. (2009). Burnout: 35 years of research and practice. *Career Development International, 14*, 204-220.

Schaufeli, W.B., & Taris, T.W. (2005). The conceptualization and measurement of burnout: common ground and worlds apart. *Work & Stress, 19*, 356-262.

Schaufeli, W.B., & Taris, T.W. (2014). A critical review of the Job Demands-Resources model: implications for improving work and health. In G. Bauer & O. Hämmig (Eds), *Bridging occupational, organizational and public health: A transdisciplinary approach*. (pp. 43-68). Dordrecht: Springer.

Schaufeli, W.B., Taris, T.W., & Van Rhenen, W. (2008). Workaholism, burnout and engagement: three of a kind or three different kinds of employee well-being. *Applied Psychology: An International Review, 57*, 173-203.

Schaufeli, W.B., & Van Dierendonck, D. (2000) *Handleiding van de Utrechtse Burnout Schaal – UBOS [Test manual: Utrecht Burnout Scale - UBOS]*. Lisse: Swets & Zeitlinger.

Schaufeli, W., Van Wijhe, C., Peeters, M., & Taris, T.W. (2011). Werkverslaving, een begrip gemeten [Work addiction: The measurement of a concept]. *Gedrag & Organisatie, 24,* 43-63.

Seidman, S.A., & Zager, J. (1987). The Teacher Burnout Scale. *Educational Research Quarterly, 11*, 26-33.

Shirom, A., & Melamed, S. (2006). A comparison of the construct validity of two burnout measures in two groups of professionals. *International Journal of Stress Management, 13,* 176–200.

Shoji, K., Cieslak, R., Smoktunowicz, E., Rogala, A., Benight, C.C., & Luszczynska, A. (2016). Associations between job burnout and self-efficacy: a meta-analysis. *Anxiety, Stress, & Coping, 29,* 367-386.

Streiner, D.L., & Cairney, J. (2007). What's under the ROC? An introduction to receiver operating characteristics curves. *The Canadian Journal of Psychiatry, 52*, 121-128.

Swider, B.W., & Zimmerman, R.D. (2010). Born to burnout: a meta-analytic path model of personality, job burnout, and work outcomes. *Journal of Vocational Behavior, 76*, 487-506.

Taris, T.W. (2006). Is there a relationship between burnout and objective performance? A critical review of 16 studies. *Work & Stress, 20,* 316–334.

Taris, T.W., Le Blanc, P.M., Schaufeli, W.B., & Schreurs, P.J.G. (2005). Are there causal relationships between the dimensions of the Maslach Burnout Inventory? A review and two longitudinal tests. *Work & Stress, 19,* 241-258.

Terluin, B. (1996). De Vierdimensionale Klachtenlijst (4DKL). Een vragenlijst voor het meten van distress, depressie, angst en somatisatie [A questionnaire for assissing distress, depression, anxiety and somatization]. *Huisarts en Wetenschap, 39,* 538-54.

Terluin B. (1998). De Vierdimensionale Klachtenlijst (4DKL) in de huisartspraktijk: psychodiagnostisch gereedschap [The Four-dimenisonal Symptom Questionnaire in general practice; a diagnistic tool]. *De Psycholoog, 33*,8-24.

Terluin, B., Van Marwijk, H., Adèr, H., de Vet, H., Penninx, B., Hermens, M., Van Boeijen, C., Van Balkom, A., Van der Klink, J., & Stalman, W. (2006). The Four-Dimensional Symptom Questionnaire (4DSQ): a validation study of a multidimensional self-report questionnaire to assess distress, depression, anxiety and somatization. *BMC Psychiatry, 6,* 30–40.

Terluin, B., Van Rhenen, W., Schaufeli, W.B., & De Haan, M. (2004). The Four-Dimensional Symptom Questionnaire (4DSQ): measuring distress in a working population. *Work & Stress, 18*, 187-207.

Van Dam, A., Keijsers, G.P.J., Eling, P.A.T.M., & Eni, S. (2012). Impaired cognitive performance and responsiveness to reward in burnout patients: two years later. *Work & Stress, 25,* 333–346.

Vanheule, S., Rosseel, Y., Vlerick, P., Van de Ven, B., & Declerc, F. (2012). De factoriële validiteit en meetinvariantie van de Utrechtse Burnout Schaal [The factorial validity and invariance of the Utrecht Burnout Scale]. *Gedrag & Organisatie, 25,* 192-201.

Van der Klink, J.T.L & Van Dijk., F.T.H. (2014). Dutch practice guidelines for managing adjustment disorders in occupational and primary health care. *Scandinavian Journal of Work, Environment & Health, 29*, 478-487.

Verbraak, M., Van de Griendt, J., & Hoogduin, K., (2006). Ontwikkeling van de BurnOut Neurasthenie KlachtenSchaal (BONKS) [Development of the Burnout Neurastenia Symptom Scale]. *Diagnostiek-wijzer, 9*, 30-49.

Verschuren, C.M., Nauta, N., Bastiaanssen, M., Terluin, B., Vendrig, A., Verbraak, M., Fliweert, S., Vreizen, J., Van Zanten-Prybysz, I., & Loo, M. (2011). *Één lijn in de eerste lijn bij overspanning en burnout: Multidisciplinaire Richtlijn Overspanning*

en Burn-out [Multidisciplairy guidelines for job stress and burnout]. Utrecht: LVE, NHG, NVAB[33]

Wheeler, D., Vassar, M., Worley, J., & Barnes, L. (2011). A reliability generalization meta-analysis of coefficient alpha for the Maslach Burnout Inventory. *Educational and Psychological Measurement, 71*, 231-244.

Widaman, K. F. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. *Applied Psychological Measurement,* 9, 1, 1-26.

Worley, J.A., Vassar, M., Wheeler, D.L., & Barnes, L.L.B. (2008), Factor structure scores from the Maslach Burnout Inventory: a review and meta-analysis of 45 exploratory and confirmatory factor-analytic studies. *Educational and Psychological Measurement, 68*, 797-823.

[33] The guidelines (in Dutch) can be downloaded from: https://www.nvab-online.nl/sites/default/files/bestanden webpaginas/MDRL_Overspanning-Burnout.pdf

Appendix 1 – The Energy Compass

Concept	Omschrijving
Job demands	
Quantitative demands	The extent to which one believes that the work cannot be done in time and / one cannot meet the required performance standards
Emotional demands	The extent to which the job affects or calls on employee's feelings.
Physical demands	The extent to which the job is physically demanding.
Bureaucracy	The extent to which rules, guidelines, regulations, procedures, and forms are seen as useless, unnecessary, restrictive or cumbersome in view of a good and efficient task performance
Role conflict	The extent to which conflicting requirements exist that prevent the employee from properly fulfilling his or her work role.
Interpersonal conflict	The extent to which work-related or personal conflicts exist with other people at work.
Harassment	The extent to which one unwanted sexual attention, threats or violence, bullying or discrimination has occurred in the past 12 months.
Work-life conflict	The extent to which the employee is having trouble in combining work and private life.
Job resources	
Role clarity	The extent to which the tasks to be performed are clearly defined and the expectations and responsibilities for the employee are clear.
Meeting expectations	The extent to which the employee does not meet the expectations of management or others.
Team atmosphere	The extent to which a pleasant atmosphere exists in the team and the employee feels at ease.
Team effectiveness	The extent to which the team collaborates effectively, the team members take their responsibility and work go the best of their abilities
Person-job fit	The extent to which the current job fits what the employee can and what he or she wants..
Opportunities for learning and development	The extent to which the employee learns new things and develops him- of herself personally as well as professionally.
Alignment	The extent to which the employee is aware of the mission of the organization and contributes to it.
Availability of tools	The extent to which the employee has sufficient resources (tools, devices, instruments) to do his or her job properly.

Engaging leadership	The extent to which the supervisor inspires, strengthens, connects, and empowers his or her followers.
Outcomes	
Organizational commitment	The extent to which the employee feels committed to the organization he or she is working for.
Team commitment	The extent to which the employee feels committed to his or her work team
Job satisfaction	The extent to which the employee is satisfied with the current job
Turnover intention	The extent to which the employee is planning to change jobs in the next year.
Workability	The extent to which the employee assesses his or her work ability, ranging from very poor to excellent
Work-related absenteeism	In how far was the sickness absence work related?
In-role performance	The extent to which the job performance of the employee meets the job requirements
Work performance	The overall assessment of the current job performance on a scale ranging from very poor (1) to outstanding (10).

Appendix 2 – Scoring forms of the BAT

Work-related version of the BAT

Instruction

The following statements are related to your work situation and how you experience this situation. Please state how often each statement applies to you.

Scoring

Never	Rarely	Sometimes	Often	Always
1	2	3	4	5

	Never	Rarely	Sometimes	Often	Always
Exhaustion					
1. At work, I feel mentally exhausted*	☐	☐	☐	☐	☐
2. Everything I do at work requires a great deal of effort	☐	☐	☐	☐	☐
3. After a day at work, I find it hard to recover my energy*	☐	☐	☐	☐	☐
4. At work, I feel physically exhausted*	☐	☐	☐	☐	☐
5. When I get up in the morning, I lack the energy to start a new day at work	☐	☐	☐	☐	☐
6. I want to be active at work, but somehow I am unable to manage	☐	☐	☐	☐	☐
7. When I exert myself at work, I quickly get tired	☐	☐	☐	☐	☐
8. At the end of my working day, I feel mentally exhausted and drained	☐	☐	☐	☐	☐
Mental distance					
9. I struggle to find any enthusiasm for my work*	☐	☐	☐	☐	☐
10. At work, I do not think much about what I am doing and I function on autopilot*	☐	☐	☐	☐	☐
11. I feel a strong aversion towards my job	☐	☐	☐	☐	☐
12. I feel indifferent about my job	☐	☐	☐	☐	☐
13. I'm cynical about what my work means to others*	☐	☐	☐	☐	☐
Cognitive impairment					
14. At work, I have trouble staying focused*	☐	☐	☐	☐	☐

15. At work I struggle to think clearly	☐	☐	☐	☐	☐
16. I'm forgetful and distracted at work	☐	☐	☐	☐	☐
17. When I'm working, I have trouble concentrating*	☐	☐	☐	☐	☐
18. I make mistakes in my work because I have my mind on other things*	☐	☐	☐	☐	☐
Emotional impairment					
19. At work, I feel unable to control my emotions*	☐	☐	☐	☐	☐
20. I do not recognize myself in the way I react emotionally at work*	☐	☐	☐	☐	☐
21. During my work I become irritable when things don't go my way	☐	☐	☐	☐	☐
22. I get upset or sad at work without knowing why	☐	☐	☐	☐	☐
23. At work I may overreact unintentionally*	☐	☐	☐	☐	☐

Note: * Short version.

	Never	**Rarely**	**Sometimes**	**Often**	**Always**
Psychological complaints					
1. I have trouble falling or staying asleep	☐	☐	☐	☐	☐
2. I tend to worry	☐	☐	☐	☐	☐
3. I feel tense and stressed	☐	☐	☐	☐	☐
4. I feel anxious and/or suffer from panic attacks	☐	☐	☐	☐	☐
5. Noise and crowds disturb me	☐	☐	☐	☐	☐
Psychosomatic complaints					
6. I suffer from palpitations or chest pain	☐	☐	☐	☐	☐
7. I suffer from stomach and/or intestinal complaints	☐	☐	☐	☐	☐
8. I suffer from headaches	☐	☐	☐	☐	☐
9. I suffer from muscle pain, for example in the neck, shoulder or back	☐	☐	☐	☐	☐
10. I often get sick	☐	☐	☐	☐	☐

General version of the BAT

Instruction

The following statements are related to your day-to-day situation and how you experience that. Please state how often each statement applies to you.

Scoring

Never	Rarely	Sometimes	Often	Always
1	2	3	4	5

		Never	Rarely	Sometimes	Often	Always
Exhaustion						
1.	I feel mentally exhausted*	☐	☐	☐	☐	☐
2.	Everything I do requires a great deal of effort	☐	☐	☐	☐	☐
3.	At the end of the day, I find it hard to recover my energy*	☐	☐	☐	☐	☐
4.	I feel physically exhausted*	☐	☐	☐	☐	☐
5.	When I get up in the morning, I lack the energy to start a new day	☐	☐	☐	☐	☐
6.	I want to be active, but somehow I am unable to manage	☐	☐	☐	☐	☐
7.	When I exert myself, I quickly get tired	☐	☐	☐	☐	☐
8.	At the end of my day, I feel mentally exhausted and drained	☐	☐	☐	☐	☐
Mental distance						
9.	I struggle to find any enthusiasm for my work*	☐	☐	☐	☐	☐
10.	I feel a strong aversion towards my job*	☐	☐	☐	☐	☐
11.	I feel indifferent about my job	☐	☐	☐	☐	☐
12.	I'm cynical about what my work means to others*	☐	☐	☐	☐	☐
Cognitive impairment						
13.	I have trouble staying focused*	☐	☐	☐	☐	☐
14.	I struggle to think clearly	☐	☐	☐	☐	☐
15.	I'm forgetful and distracted	☐	☐	☐	☐	☐
16.	I have trouble concentrating*	☐	☐	☐	☐	☐
17.	I make mistakes because I have my mind on other things*	☐	☐	☐	☐	☐

Emotional impairment						
18.	I feel unable to control my emotions*	☐	☐	☐	☐	☐
19.	I do not recognize myself in the way I react emotionally*	☐	☐	☐	☐	☐
20.	I become irritable when things don't go my way	☐	☐	☐	☐	☐
21.	I get upset or sad without knowing why	☐	☐	☐	☐	☐
22.	I may overreact unintentionally*	☐	☐	☐	☐	☐

Note: * Short version.