



Validation of a Text-based Measure of Job Affect

Development of a Context-Specific Computer-Aided Text Analysis Tool

ERASMUS UNIVERSITY ROTTERDAM

ERASMUS SCHOOL OF ECONOMICS

Master Thesis Behavioural Economics

Presented by: Tijmen Mulder

Student ID number: 413834

Supervisor: Indy Wijngaards

Second assessor:

Date final version: August 7, 2019

Abstract

This study investigates how to construct a valid text-based measure of the job affect dimensions boredom and enthusiasm using computer-aided text-analysis (CATA). Following a recent guideline for CATA research (McKenny et al., 2018; Short et al., 2010), the study aim was to conduct a systematic approach towards validation of the text-based measures and analyse the convergent and discriminant validity to job- and life-related behaviour constructs.

The text-based measures were developed from responses to a semi-open question in a job satisfaction survey with 364 English-speaking respondents. Interrator reliability tests resulted in scores that were indicative of high reliability. Convergence validity results showed that the convergent correlation of the text-based measure of boredom was significantly stronger than discriminant correlations. The convergent correlation of the text-based measure of enthusiasm, however, was not significantly stronger than discriminant correlations. The relationships offered substantial evidence for the discriminant validity of the constructed text-based measures.

In the end, the limitations of this study were discussed that could be addressed in future research. The study showed that semi-open questions with job affect words can be used to construct a context-specific text-based measure to analyse the relationships with job- and life-related behaviour constructs. Therefore, it might be interesting to expand this study field in further CATA research by analysing the added value of open questions in similar survey instruments which considered closed questions as the standard. The study could also function as an inspiration for scholars interested in content analysis studies of job satisfaction using CATA.

Keywords: CATA, job affect, open questions, semi-open question, text-based measure, reliability, convergence validity, discriminant validity

Content

Abstract	2
1. Introduction	4
2. Theoretical Framework.....	7
2.1 Added Value of Open Questions and CATA	7
2.2 Construct Validity	8
2.2.1 Reliability	8
2.2.2 Convergent Validity	9
2.2.3 Discriminant Validity	10
3. Method	13
3.1 Participants and Data Collection	13
3.2 Measures	13
3.3 Construction of Text-based Measure	14
3.3.1 Deductive Analysis	15
3.3.2 Inductive Analysis.....	16
3.4 Data Processes and Statistical Analysis	18
4. Results	19
4.1 Reliability	19
4.2 Convergent and Discriminant Validity.....	19
5. Discussion and Conclusion.....	22
5.1 Limitations and Further Research	23
References.....	25

1. Introduction

At a time in which many employees are becoming victim of burnouts, and their workload are becoming more stressful (CBS, 2018), measuring job satisfaction is an important task for managers. The significant amount of research that has been conducted in this area supports this statement (see e.g. Aziri, 2011; Judge et al., 2017; Spector, 1997). Research has shown that job satisfaction has an impact on an employee's motivation, and thereby on someone's productivity and, eventually, the overall performance of an organisation. Job affect is conceptually related to job satisfaction (Connolly & Viswesvaran, 2000), and is an essential component of job attitude and potentially a significant predictor of certain job behaviours (George, 1989; Weiss and Cropanzano, 1996), such as personal-organisation fit (P-O fit, Herrbach & Mignonac, 2006), turnover intention (Carsten & Spector, 1987; Hom, Katerberg & Hulin, 1979; Hulin, 1966; 1968; Mobley, Horner, & Hollingsworth, 1978), and organisational citizenship behaviour (OCB, Bateman & Organ, 1983; Farrell, 1983). In the past, the focus of job satisfaction questionnaires was placed on cognitive aspects and the affective antecedents were generally ignored (Chaufeli, 2006; Organ & Near, 1985). Today, in response to the earlier mentioned findings, we have seen a switch to a more affective approach with as a result a growing interest in related studies by researchers (Agho, Price & Mueller, 1992; Bateman & Organ, 1983; Connolly & Viswesvaran, C., 2000; Weiss, 2002).

As job satisfaction is a complex but important concept, various survey question instruments have been developed to measure job satisfaction (Kaplan et al., 2009; Van Saane et al., 2003) e.g., the Job description index (Weiss, Dawis & England, 1967) and the Minnesota satisfaction questionnaire (Smith, 1969). The importance of the measurement method used to measure job satisfaction should not be underestimated, as only a few methods have shown both high reliability and high validity (Van Saane et al., 2003). Most of the job affect items in these survey methods use closed question instruments, while open question instruments have rarely been used in job satisfaction constructs (Gilles et al., 2017). One of the main advantages of using closed questions is that they are relatively easy to collect, code and adapt to existing model frameworks (Glerum, Atasoy & Bierlaire, 2013). Closed questions can be easier analysed compared to open questions and are easy to interpret (Glerum, Atasoy & Bierlaire, 2014), however, Looker et al. (1989), point out that open questions are undervalued in most surveys and often only used to complement closed questions. Open questions can be used to overcome measure limitations of closed questions as they permit the retrieval of constructs related to the content of language, e.g. emotions or affection (Petroni & Whissel, 1988). Moreover, the

diversity of the respondent's answers is more extensive in open questions than in closed questions (Schuman & Presser, 1996). Therefore, it can be useful to increase the number of open questions used in a survey when measuring job affect or other related constructs in job satisfaction surveys. The major issue with open questions, however, is that it is more challenging to construct text-based measures because of the need to use human coding and validation (Zehner et al., 2016) to obtain good results.

Computer-aided text analysis (CATA) can help solve these limitations of analysing open questions. Weber (1990) describes in his book how content analysis can be used to help to make text data more relevant and understandable using a set of procedures to make valid inferences from the text. The basic idea is that the main text is divided into several content categories, consisting of one or more words based on the frequency of those words. Each category represents an overall meaning related to the words in that specific group. The focus of this study was to construct text-based measures using CATA with affective words related to job satisfaction as a unit of measurement. The dataset used in this content analysis was a survey sample consisting of 364 employees answering a semi-open job satisfaction item. The semi-open question was: Which three to five adjectives come to mind when thinking of your job as a whole?

Two words of two different aspects of the emotion scale, boredom and enthusiasm, were analysed to measure their convergence with closed question items of the job-related affective well-being scale (JAWS, Van Katwyk, Fox, Spector & Kelloway, 2000), which was included in the same survey as the semi-open question item. The JAWS is a closed question scale that can be used to assess employee's emotional reactions to a specific task or to their job in general. Research has proved the effects boredom (Bench & Lench, 2013; Tsai, 2016) and enthusiasm (Wood, 1998) can have on job motivation and other behaviours. Moreover, the emotions boredom and enthusiasm may have connections with other related constructs to job satisfaction (e.g. Halbesleben & Wheeler, 2008; Kristof, 1996; Schaufeli & Salanova, 2014).

The scientific relevance of this research is twofold. First, the study should provide insights into the measurement quality of a semi-open question for measuring job affect. The scientific knowledge on this topic is rather limited, however, there are clear potentials and advantages of using semi-open questions and text-based measures to measure employee behaviour (Borg & Zuell, 2012). Analysing the convergence validity between a semi-open affect question item and related constructs will add to the body of knowledge. Second, CATA has not previously been

used in studies related to job affect, even though CATA is more systematic in its validation and research has proved the role of affectivity in job satisfaction (George, 1989; Weiss & Cropanzano, 1996). Consequently, this study contributes to existing literature by providing an analysis of the discriminant validity of the constructed text-measures compared to other closed question job satisfaction measures. The practical relevance of the study could be an increase in the use of open question items in employee surveys to improve the quality of survey responses.

Following a recent guideline for computer-aided text analysis research (McKenny et al., 2018; Short et al., 2010), four study aims were determined. The first two objectives were to construct the text-based measures and enhance reliability in this process. The third objective was to analyse the convergent validity of the constructed text-based measures with the JAWS. The final objective was to examine the discriminant validity of the text-based measures compared to other job- and life-related behaviour constructs, e.g. personal-organisation fit, job satisfaction, turnover intention, organisational citizenship behaviour and life satisfaction.

2. Theoretical Framework

In this section, a short outline of the added value of open questions in measuring job affect will be provided, followed by a discussion of the increased convenience of constructing text-based measures and the working and added value of CATA. Finally, a theoretical framework will be described based on guidelines which are designed to promote reliability (McKenny et al., 2018) and validity (Short et al., 2010). In this way, the theoretical framework should substantiate the stated hypotheses and help answer the research objectives. The discussion on the development of the dictionaries related to this study is given in the methodology section.

2.1 Added Value of Open Questions and CATA

Most of the surveys designed to measure job satisfaction use closed questions instead of open questions (Gilles et al., 2017), with only open questions such as: “Please give us feedback or comments about your job” being used. Even though the study of Gilles et al. (2017) is one of the few studies that analysed the added value of such questions is it questionable in which manner it correlates with closed job satisfaction questions. In general, the construction of text-based measures regarding open questions has been difficult and is often avoided. The main disadvantages of this construction are the reliance on human coding (Wolfe et al., 1993) and the possibilities to enhance the validity of the text-based measure (Zehner et., 2016). The limited amount of research on text-based measures done to date, i.e. in job satisfaction studies, is mostly due to these drawbacks (Short et al., 2010). Fortunately, several content analyses software packages have been developed, which can be helpful when constructing measures based on text such as open questions (Duriau, Reger & Pfarrer, 2007). Consequently, the convenience of constructing text-based measures has increased, also for measures related to job satisfaction (Borg & Zuell, 2012; Gilles et al., 2017).

One type of analysis used to support this coding process is CATA, defined by Wolfe, Gephart and Johnson (1993) as follows: “CATA is constituted by software programs that facilitate the analysis of textual data” (p.638), CATA makes it possible to analyse textual data with higher reliability than human coding, and it can do this faster, and with lower costs (Neuendorf, 2002; Wolfe, Gephart & Johnson, 1993), than humans. A meta-study by Duriau and colleagues (2007) showed that even though there have been many studies on the use and advantages of CATA, only 25% of the articles reviewed (24 out of 98) used this analysis tool. One possible reason for this limited use could be that it is still difficult to deal with validity and implement such an analysis in an empirical research situation. Besides, it is more challenging to capture contextual factors (Kippendorff, 2004). Short et al. (2010) and McKenny

et al. (2018) presented two studies, designed to help researchers develop content analyses with both higher reliability and validation. Their work was used as a guideline in the following section.

2.2 Construct Validity

2.2.1 Reliability

When constructing a text-based measure, it is important to enhance the reliability of the measure (McKenny et al., 2018). According to Carmines and Zeller (1979), the reliability of an estimator refers to the extent to which an experiment, test, or any measuring tool gets the same result in repeated analyses (p.11). Estimators always come with a certain level of measurement error, creating the goal for every researcher to minimise this as much as possible. Neuendorf (2002) found more reliability when using CATA and showed that it is not always necessary to require human coding. Short et al. (2010) showed that CATA is more accurate, capable of analysing larger datasets with high speed and is less costly than human coding. These advantages have led to a number of studies using it in various constructs (Short et al., 2010; McKenny et al., 2018), however, with the use of CATA may also carry with it certain measurement errors, and therefore, it is still important to consider the reliability of this method.

Schmidt et al. (2003) and McKenny et al. (2018) found three sources of measurement error when using CATA: transient error, specific factor error and algorithm error. First, transient error due to persistent temporal factors, e.g. mood, which might have an affect on responders' answers, and transient error may increase by processes over time. The size of the transient error can be measured using a test-retest estimate, which tests the correlation between texts over time and thereby its consistency (Schmidt & Hunter, 1996). Despite its relevance, it was not possible to use this test-retest estimate for the research reported here because the dataset did not consist of responses collected at two points in time. Second is the specific factor error, which is likely to be an existing measurement error when using CATA, and which will threaten both validity and reliability. A specific factor error arises when wording of the question of a survey influence the answers of respondents (Schmidt & Hunter, 1999). Specific word choices in the content of measure may influence the survey scores. Relating this specific error to generating dictionaries used by CATA we have to look at the content of word lists. Developing new dictionaries by selecting new specific words considered relevant to the survey environment may increase the specific factor error. Therefore, it is essential to measure this error using a reliability test that measures the extent to which human coding and CATA coding produces similar scores. In this study, this test was done using Holsti's (1969) method

for assessing interrater reliability. Third, the algorithm error, which arises when two CATA software programs produce different scores when analysing the same text: CATA is subject to this type of measurement error, which is not relevant in traditional measurement approaches. It is possible to find the size of this error by analysing the extent to which two CATA software packages are similar to each other. Krippendorff's alpha is a useful test for this (Krippendorff, 2018). The algorithm error, however, was not possible to test as there were no free software available besides CAT Scanner.

While the specific factor error can be determined, the transient error and algorithm error can not be determined or estimated and are therefore not included in the methodology and result sections. Consequently, it is hypothesized that:

Hypothesis 1: The CATA and human coding produce similar scores (Specific factor error).

2.2.2 Convergent Validity

Besides ensuring the reliability of using CATA, it is essential to enhance validation when developing the dictionaries to be used for the text-based measures. This can be done using both a deductive and an inductive approach. As suggested by Short et al. (2010), it is advisable to start with a deductive approach to CATA, especially in the case of a text-based measure that is being developed for a new concept, as was applicable for this study. A deductive approach is theory-driven, and, if possible, allowed to be used and extend in other related research fields (Joffe & Yardley, 2004). Following the recommended procedures of Short (2010), there are four steps that need to be taken to enhance deductive content validity: (1) create a working definition of the construct of interest, (2) make an initial assessment of construct dimensionality based on existing literature, (3) develop an exhaustive list of keywords from the formal definition to capture the construct of interest and (4) validate word lists using content experts and assess reliability.

In addition, an inductive approach is provided, which is rather data-driven. Short et al. (2010) believe in the contribution of this approach as it can complete the deductive approach by following a different procedure in the process of finding potential words related to the unit of measurement. Following the recommended procedures of Short (2010) again, there are five steps that need to be taken to enhance inductive content validity: (1) identify commonly used words from narrative text of interest using CATA software, (2) identify or create a working definition of the construct of interest to guide word selection, (3) identify words that match

the construct of interest, (4) establish initial interrator reliability and (5) refine and finalize word lists.

Combining a deductive and inductive approach seems to be the leading procedure and fills the gap between theory and practice driven characteristics. The combination has been used more often in content analysis research and has proven to be the best approach (Fereday & Muir-Cochrane, 2006; Kabanoff, Walderssee, & Cohen, 1995). Therefore is assumed that a combined strategy with both a deductive and an inductive expanding of the dictionaries will result in a more extensive word list related to boredom and enthusiasm than using only a deductive approach.

In this study, the convergent validity of the constructed text-based measures of both boredom and enthusiasm were analysed against another instrument, the JAWS, that is measuring the same construct in closed question items. It was expected that this relationship is relatively stronger compared to other related constructs. The above gives rise to the following hypotheses:

Hypothesis 2a: The expected relationship between the text-based measure of boredom and the JAWS boredom is positive and stronger than with other related constructs.

Hypothesis 2b: The expected relationship between the text-based measure of enthusiasm and the JAWS enthusiasm is positive and stronger than with other related constructs.

2.2.3 Discriminant Validity

Besides analysing the convergence between the semi-open job affect question and the JAWS, it is also relevant to look to instruments that measured a related, but different construct. The criterion for a sufficient degree of discriminant validity is determined at a correlation of 0.50 or less (Van Saane, 2003). The convergence was tested following existing comprehensive networks (Bowling & Hammond, 2008). One antecedent, i.e. person-organisation fit, one correlation, job satisfaction, and three outcomes, i.e. turnover intention, organisational citizenship behaviour and life satisfaction, were considered.

The first construct, P-O fit, concerns the degree an employee's values match or fit those of the organisation (Kristof-Brown, 2005) and has a strong impact on outcomes, e.g. OCB and turnover intention (Sjöberg & Sverke, 2000; Yaniv, Lavi & Siti, 2010). Employees become satisfied when their personal values fit with those of the job (Kristof, 1996). Therefore it is

expected that the correlation of enthusiasm and boredom will be positive and negative, respectively. Hence:

Hypothesis 3a: The text-based boredom measure correlates negatively with P-O fit.

Hypothesis 3b: The text-based enthusiasm measure correlates positively with P-O fit.

The second construct that was analysed was job satisfaction. Conceptually, job satisfaction is closest to job affect, and the role of affectivity in job satisfaction has been proven for both positive and negative affects (Agho, Price & Mueller, 1992; Bateman & Organ, 1983; Connolly & Viswesvaran, C., 2000; Judge, 2017; Weiss, 2002). As described, enthusiasm is high pleasurable and boredom low pleasurable (Van Katwyk et al., 2000). Therefore:

Hypothesis 3c: The text-based boredom measure correlates negatively with job satisfaction.

Hypothesis 3d: The text-based enthusiasm measure correlates positively with job satisfaction.

Another construct that is related to job affect is turnover intention, a person's willingness to change jobs or companies (Tett & Meyer, 1993). Enthusiasm is a component of work engagement (Russel, 1980; Schaufeli & Bakker, 2004), which is negatively related to turnover intention (Halbesleben & Wheeler, 2008). Therefore a negative correlation is expected between the text-based enthusiasm measure and turnover intention. The consequences of boredom stimulate turnover intention (Schaufeli & Salanova, 2014). Thus:

Hypothesis 3e: The text-based boredom measure correlates positively with turnover intention.

Hypothesis 3f: The text-based enthusiasm measure correlates negatively with turnover intention.

The fourth construct, OCB, concerns voluntary help or assistance of employees within the organisation without any promise of receiving a reward for the behaviour (Organ, 1988).

Positive affectivity, such as enthusiasm, is expected to increase OCB, while negative emotion is expected to decrease OCB (Bateman & Organ, 1983; Farrell, 1983). Therefore:

Hypothesis 3g: The text-based boredom measure correlates negatively with OCB.

Hypothesis 3h: The text-based enthusiasm measure correlates positively with OCB.

The final construct tested was life satisfaction. Life satisfaction refers to a judgemental process in which individuals, based on their own criteria, value the quality of their life (Shin & Johnson, 1978). It is therefore not as much related to job affect as the earlier mentioned constructs. Nevertheless, many studies have proven the influence job satisfaction, and indirectly job affect, has on people's life (see e.g. Heller, Judge & Watson, 2002; Judge & Watanabe, 1993). Since this indirect effect exists is a positive correlation expected for enthusiasm and a negative correlation expected for boredom. Thus:

Hypothesis 3i: The text-based boredom measure correlates negatively with life satisfaction.

Hypothesis 3j: The text-based enthusiasm measure correlates positively with life satisfaction.

3. Method

3.1 Participants and Data Collection

A virtual crowdsourcing platform, Prolific Academic, was used to collect the data used in this study. Peer and colleagues (2017) found participants were more naïve and honest using Prolific in comparison to another often used platform such as Amazon Mechanical Turk. The employees in this survey all had an American nationality, and their responses were guaranteed to be anonymous. Whenever a respondent had filled in the survey too fast, which was possible to determine, the responses were not used in the final dataset to prevent input from any unthinking responders. The survey administration software used for the data was qualtrics.

All the respondents worked at least ten hours a week, almost 57% of the employees 40 hours with an average of 37 hours. The final data collection consisted of 364 valid responses. Most of the employees were male (55.8%), and the average age of the respondents was 35.1. Almost all of the respondents had been to college (93.8%) and the majority had a permanent contract (76.7%). The average number of years of experience in their job was 5.4.

3.2 Measures

Two text-based measures based on the answers respondents gave in a semi-open question were developed to analyse convergence validity and discriminant validity with other related constructs. The data measures used in this study are outlined below.

Job affect. The JAWS is a closed item scale developed to measure an individual's emotional responses to their job (Van Katwyk et al., 2000). In this study, the scale exists of 20-items, and the responses used a five-point response scale on which people could indicate how often they experienced a related emotion. The anchors on this scale were *Never, Rarely, Sometimes, Quite often and Always*. The 20-items could have either a positive or negative emotional value (valence), while these also could differ in the degree of positivity or negativity, i.e. activation. The complete list of emotion could be divided into four parts, categorized by two often-used dimensions within studies related to emotions: pleasure and arousal (Russel, 1980). One: positive emotions with a high level of activation, e.g. enthusiastic and energetic. Two: positive emotions with a low level of activation, e.g. calm and satisfied. Three: negative emotions with a high level of activation, e.g. angry and disgusted. Four: negative emotions with a low level of activation, e.g. bored and depressed. As earlier mentioned the focus of this study was on the boredom- and enthusiasm-items of the scale.

P-O fit. The P-O fit was developed by Cable and Judge (1996) and is measured using a 3-item scale. The answer categories range from 1 (*not at all*) to 5 (*completely*). An example question is “Do you think the values and “personality” of this organisation reflect your own values and personality?”

Job Satisfaction. Job satisfaction was measured using two multiple-item instruments, the Michigan Organisational Assessment Questionnaire Job Satisfaction Subscale (MOAQ-JSS, Cammann et al., 1979) and the Brief Index of Affective Job Satisfaction (BIAJS, Thompson and Phua, 2012). Both instruments focus on the affective components of job satisfaction. The answer categories on the MOAQ-JSS ranges from 1 (*strongly disagree*) to 7 (*strongly agree*). One of the questions states, “In general, I like working here”. As this instrument includes a reverse-code item does it control for careless responding. The answer categories on the BIAJS ranges from 1 (*strongly disagree*) to 5 (*strongly agree*). An example question is “My job is unusual.”

Turnover Intention. Turnover intention was measured using the three-item turnover intention subscale in the MOAQ (Cammann et al., 1979). The scale has answer categories ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). It has one reverse-code item with answers ranging from 1 (*extremely likely*) to 5 (*not likely*) to prevent careless responding. One of the questions states, “I will probably look for a new job in the next year.”

OCB. OCB was measured using the 10-item Organisational Citizenship Behaviour Checklist (OCB-C) by Spector et al. (2010). The answer categories of this scale range from 1 (never) to 5 (every day). An example question is “Volunteered to attend meetings or work on committees on own time.”

Life satisfaction. Life satisfaction was measured using a 5-item scale called the Satisfaction With Life Scale (SWLS, Diener et al., 1985). The answer category ranges from 1 (strongly disagree) to 7 (strongly agree). An example question is “So far I have gotten the important things I want in life.”

3.3 Construction of Text-based Measure

In this section is the procedure described for generating dictionaries of the particular job affects boredom and enthusiasm. There were no pre-existing dictionaries that were suitable for the purpose of this study, therefore it was necessary to develop a customised dictionary. Following Short et al. (2010), this was done using both a deductive and an inductive approach, and taking assessing reliability into consideration (McKenny et al., 2018). As boredom and enthusiasm in this analysis were specifically focused on job behaviour they were not conceptualised as multidimensional constructs. Two single word lists were created using the

following steps. The data for the inductive approach was taken from the responses to the semi-open job satisfaction survey question “Which three to five adjectives come to mind when you think about your job as a whole?” The generated wordlists were then used in CAT Scanner (Mckenny, Short & Newman, 2012) to create the two text-based measures.

3.3.1 Deductive Analysis

A deductive or theory-driven approach was used as the basis for developing the dictionaries using the work of Short et al. (2010) as a baseline.

The first step for constructing the dictionaries was to identify a formal definition of the concept of interest. The concept of interest in this research was job affect and for this study the units of analysis were the keywords boredom and enthusiasm. There is no one universal definition of boredom, but as the focus of this study was job affect, a more work-related boredom definition was used for deductive development. The following definition was used: “A negative, affective motivational state of mind that results from a work environment that elicits little activity and provides little challenge and pleasure” (Reijseger et al., 2013, p. 509). As earlier mentioned is this emotion been categorized as dissatisfaction and low arousal. In contrast to boredom is enthusiasm, a more satisfied, high arousal emotional state. Relating enthusiasm to work-engagement, Stevenson and Waite (2011) state enthusiasm can be described as: “An intense and eager enjoyment, interest, or approval towards a job-related environment.”

The second step involved an initial assessment of construct dimensionality based on existing literature. As there were no such dimensions found, as seen in other CATA studies which covered larger concepts, e.g. Entrepreneurial Orientation (Short et al., 2010), this step for both boredom and enthusiasm was not applicable in this study.

In the third step, an extensive wordlist was created to find all variants and synonyms of boredom and enthusiasm. Rodale’s (1978) The Synonym Finder was used in combination with an online thesaurus to generate emotion-specific dictionaries. These emotions should fit the context of the structure and they were chosen carefully to avoid any misplacement. The main focus was on finding adjectives as adjectives were expected to capture the word responses of the semi-open question. Some nouns and verbs were included, as sometimes respondents used different word categories.

The fourth step for both keywords was to identify whether the words fit the construct or not. This was done in combination with an interrater reliability test using Holsti’s (1969) model, which is accepted and outlined in the result section. The final deductive dictionaries for boredom and enthusiasm consist of 101 and 107 words, respectively, as shown in Table 1.

3.3.2 Inductive Analysis

The CATA software program CAT Scanner was used for the inductive or data-driven approach; CAT Scanner inductively generates word lists from the responses on the semi-open question using words that are used three or more times. The following step was to assign these words to the boredom and enthusiasm deductive generated dictionaries manually, however, CAT Scanner did not generate many new words compared to the already deductively generated words. The main reason for this was the relatively small dataset in this study as it only consisted of 364 participants who had to give three to five adjectives in the related semi-open question. The presence of an already developed deductive dictionary captured most of the final dictionaries. Nevertheless, the inductively generated words were still an addition for the complete dictionary adding nine words to the boredom dictionary and five words to the enthusiasm dictionary. The total amount of words for boredom and enthusiasm were 110 and 112, respectively, and are shown in Table 1.

Table 1.

Complete dictionaries of keywords boredom and enthusiasm

Job Affect Dimension	Content Analysis Words
Boredom	apathy, arid, banal, blah, bland, blasé, bore, boreal, bored, boresome, boring, colourless, commonplace, dead, deadly, deadness, disinterest, dispassion, doldrums, drab, dreariness, dreary, drudging, dry, dull, dullness, ennui, flagging, flat, flatness, ho-hum, humdrum, impassivity, indifference, indifferent, insentience, insentient, insipid, irksome, jaded, jading, jejune, lacklustre, lame, leaden, lethargy, lifeless, long, long-drawn, long-winded, mind-numbing, monochromatic, monotone, monotonous, monotony, mundane, old, ordinary, pedestrian, pointless, ponderous, prosaic, reaming, repetitious, repetitive, repetition, routine, run-of-the-mill, schedule, sleepy, slow, soporific, spiritless, stale, standard, stereotyped, stodgy, stuffy, tame, tedious, tedium, tiresome, tired, tiring, trite, unanimated, uneventful, unexciting, uninspired, uninspiring, uninterested, uninterestedness, uninteresting, unvaried, vapid, weariful, wearing, wearisome, weary, wearying, wordy
Additional inductively derived words	annoying, calm, decent, easy, quiet, relaxed, relaxing, safe, stress-free
Enthusiasm	active, agog, all-fired, ambitious, animated, animatedly, animation, ardent, ardently, ardour, avid, awesome, breathless, buoyancy, buoyant, burning, cheerful, cool, cordial, crazy, eager, eagerly, eagerness, ebullience, ebullient, ebulliently, ecstatic, effervescent, effervescently, effusive, effusively, energetic, energetically, engage, engagement, enthuse, enthused, enthuses, enthusiastic, enthusiastically, excitable, excited, excitement, exciting, exuberance, exuberant, exuberantly, fanatic, fanatical, fanatically, fanaticism, fervency, fervent, fervently, fervid, fervidly, fervour, feverish, gaga, geek, glowing, great, gung-ho, gushing, gushingly, hearty, hopped-up, hot, intense, intensity, intentness, interesting, keen, keenly, keenness, life, liveliness, lively, mad, magnificent, nuts, optimism, passion, passionate, passionately, possessed, pumped, rabid, rabidly, rah-rah, rapturous, thrilled, vehement, vehemently, vital, vitality, vivacious, vivacity, warm, warmth, wholeheartedly, wholehearted, wild, willing, zeal, zealous, zealously
Additional inductively derived words	enjoyable, entertaining, fun, happy, positive

Source: Deductive word lists are developed with the aid of Rodale's (1978) *The Synonym Finder and Power Thesaurus*. The additional inductive words are generated by CAT Scanner.

3.4 Data Processes and Statistical Analysis

Data pre-processing was done in R (R Core Team 2014), CAT scanner was used to generate the inductive dictionary and to create the CATA measures, statistical analysis and tests were done in Stata (StataCorp, 2017).

In general, using CATA in context analysis should enhance a higher level of reliability compared to human coding (Short et al., 2010). The source of measurement error that is common to CATA studies and was possible to test in this thesis was the specific factor error. To test hypothesis 1, an interrater reliability test was conducted to determine if human scores were similar to the words of the generated dictionaries. At least 10% of the generated words were rated by three independent individuals to identify whether the words were representative of the identified job affect. These words were chosen randomly from both the dictionaries of boredom and enthusiasm. The specific question that the “judges” were asked was: “Would this word be representative of the construct if it was present in the related dictionary?” Although there is no generally accepted minimum established for interrater reliability, Krippendorf (2012) suggest values greater than .8 are indicative of high reliability.

Hypothesis 2a and 2b were tested by results of the computed bivariate correlations and Steiger’s (1980) z test. The former gives an indication of the direction and how strong the relationship is between the text-based measures and the JAWS instrument. The JAWS instrument has been used as it measures the same construct as the text-based measures in closed question items. Steiger’s z-test shows how strong the relationship is between the text-based measures and the JAWS instrument, compared to the relationships the text-based measures has with the other constructs. Based on these results, it is possible to answer the stated hypotheses in relation to the convergence validity.

The final hypotheses (3a-3j) are answered by the same computed bivariate correlations as discussed for the convergent validity. Based on the direction between the text-based measures and the related constructs is it possible to test and discuss discriminant validity.

4. Results

The developed dictionaries were based on Short et al.'s (2010) and McKenny et al.'s (2018) recommendations to enhance a systematic approach towards validation of the text-based measures. The results are described and answers are provided to the earlier stated hypotheses, which are focused on construct validation issues in CATA: reliability, convergent validity and discriminant validity.

4.1 Reliability

An interrater reliability test was conducted to determine if human scores were similar to the words of the generated dictionaries. The overall interrater reliability of the boredom dictionary after this test was .91 and the interrater reliability of the enthusiasm dictionary was 1. Although there is no generally accepted minimum established for interrater reliability, Krippendorff (2012) suggest values greater than .8 are indicative of high reliability and as this is the case does it support hypothesis 1.

4.2 Convergent and Discriminant Validity

An inductive approach to generate additional words for the dictionaries for boredom and enthusiasm was used. A deductive approach, more theory driven, formed the basis of the dictionary. The inductive approach generated several additional words for the dictionaries following data analysis of the responses to the semi-open question. In the dictionary of boredom, the inductive approach added another nine words and in the dictionary of enthusiasm five words. This demonstrated the additional value of doing both approaches instead of only one approach.

The convergent validity of the text-based measures was assessed by computing Pearson bivariate correlations (Table 2) and using Steiger's z test. The convergent validity of the constructed text-based measures for boredom and enthusiasm was analysed against another instrument, the JAWS, which measures the same construct in closed question items. As these instruments seem to be very related, the correlation was expected to be close to 1. The results of the computed bivariate correlations are shown in table 2. As expected was the relationship between the text-based measure of boredom and the JAWS measure of boredom strong ($r = .507$). The relationship between the text-based measure of enthusiasm and the JAWS measure of enthusiasm was slightly weaker ($r = .397$). The results also showed that, as expected, the text-based measure of boredom had a positive correlation with the JAWS boredom and a negative correlation with the JAWS enthusiasm. This was the same for the text-based measure

of enthusiasm which positively correlated with the JAWS enthusiasm and negatively correlated with the JAWS boredom. Steiger z test indicated that the text-based measure of boredom correlated significantly stronger ($z = 11.02, p < .01$) with the JAWS boredom measure than with the third and fourth highest correlating measures, job satisfaction and the JAWS enthusiasm. However, Steiger z test did not indicate that the text-based measure of enthusiasm correlated significantly stronger ($z = -0.247, p = .402$) with the JAWS enthusiasm measure than with the third and fourth highest correlating measures, job satisfaction and turnover intention. These results support hypotheses 2a and showed that the convergent correlation of the text-based measure of boredom was significantly stronger than discriminant correlations. Hypothesis 2b is not supported, as the convergent correlation of the text-based measure of enthusiasm was not significantly stronger than discriminant correlations.

Table 2.

Means, Standard Deviation and Pearson correlations of study variables (N=364)

Variable	Mean	SD	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. CATA Boredom	0.632	.937									
2. CATA Enthusiasm	0.316	.532	<i>-.152</i>								
3. JAWS Boredom	2.701	1.195	.507	<i>-.209</i>							
4. JAWS Enthusiasm	3.066	1.165	<i>-.382</i>	.397	<i>-.430</i>						
5. P-O fit	3.329	1.021	<i>-.359</i>	.297	<i>-.434</i>	.611					
6. Job satisfaction	3.646	1.174	<i>-.370</i>	.415	<i>-.526</i>	.711	.691				
7. Turnover intention	2.781	1.321	.326	<i>-.317</i>	.459	<i>-.564</i>	<i>-.541</i>	<i>-.751</i>			
8. OCB	2.778	.783	<i>-.288</i>	.078	<i>-.279</i>	.344	.314	.262	<i>-.162</i>		
9. Life satisfaction	3.137	1.113	<i>-.234</i>	.243	<i>-.263</i>	.451	.442	.553	<i>-.438</i>	.158	

Note. All correlations are significant at the level of $p < .05$ (2-tailed), except for the ones in italic. CATA = Computer-aided textual analysis; JAWS = Job-related Affective Well-being Scale; P-O = Person-organisation; OCB = Organisational citizenship behaviour; SD = Standard deviation.

The correlation matrix shown in Table 2 was also used to assess the discriminant validity of the text-based measures. Focusing on the text-based measure of boredom, the relationship is negative to the antecedents (P-O fit), correlation (job satisfaction) and outcomes (OCB and life

satisfaction). Boredom is positively related to the outcome turnover intention ($r = .326$). The strongest correlation is found with the construct of job satisfaction ($r = -.370$) and the weakest with life satisfaction ($r = -.234$).

The text-based measure of enthusiasm is positively related to the antecedents (P-O fit), correlation (Job satisfaction) and outcome (Life satisfaction). Enthusiasm is negatively related to the outcome turnover intention ($r = -.317$). The strongest correlation is found with the construct of job satisfaction ($r = .415$) and the measure did not correlate with OCB ($r = .078$). The described relationships to the related construct do offer substantial evidence for the discriminant validity of the constructed text-based measures and support the hypotheses (3a-3j).

5. Discussion and Conclusion

The purpose of this study was to construct a text-based measure of job affect. Following a recent guideline for CATA research (McKenny et al., 2018; Short et al., 2010), the study aim was to conduct a systematic approach towards validation of the text-based measures and analyse the convergent and discriminant validity to related job- and life-related behaviour constructs. The text-based measure was constructed from responses to a semi-open question in a job satisfaction survey with 364 respondents. The focus was on two different aspects of the emotion scale, boredom and enthusiasm, and the convergence and discriminant validity of the two text-based measures used with related constructs, e.g. personal-organisation fit, job satisfaction, turnover intention, organisational citizenship behaviour and life satisfaction. The systematic approach in generating the dictionaries that has been used in the development of valid and reliable text-based measures was one of the main contributions in this study.

One of the objectives was to construct the text-based measures. Two different dictionaries were generated to describe the job affects boredom and enthusiasm. As there were no existing dictionaries available for these words, it was necessary to develop valid custom dictionaries. A deductive and an inductive approach were used to develop these dictionaries in a more systematic way than often is done in similar CATA studies. Generating words using both approaches resulted in more extensive dictionaries which demonstrate the additional value of doing both approaches instead of only one approach.

The reliability of the text-based measures was checked following an interrater reliability test for the specific factor error and the results showed values that were generally established as indicative of high reliability. The JAWS instrument was used to analyse the convergence validity and showed relatively high scores with the text-based measures of boredom and enthusiasm. The relationship was not as close to 1 as expected but the text-based measure of boredom especially, still, had a sufficient level of convergence validity. There could be several explanations for why the correlations were not as close to a correlation of 1 as expected. The most important explanation might be that responders were allowed to fill in every random adjective that came to their mind. The answers could be diverse and not fully focused on one specific job affect. For example, if someone fills in “satisfied, happy and focused”, the text-based measure of enthusiasm scores it relatively low as only one word is included in the enthusiasm dictionary. On the JAWS instrument of enthusiasm, the same responder could get a high score, and so the correlation between the text-based measure and the JAWS of enthusiasm was not that strong.

Results of Steiger's z test showed that convergent correlation of the text-based measure of boredom was significantly stronger than discriminant correlations. The convergent correlation of the text-based measure of enthusiasm, however, was not significantly stronger than discriminant correlations. The discriminant validity was analysed by the bivariate correlations between the text-based measures and other job- and life-related constructs. The relationships offered substantial evidence for the discriminant validity of the constructed text-based measures. Both text-based measures had a relatively strong relationship with job satisfaction. This supports the suggested theoretical relationship between the two constructs and could prove that semi-open questions can be used to gain information for specific job behaviour and satisfaction.

5.1 Limitations and Further Research

There are some limitations in this study that could be addressed in future research concerning: (1) the reliability, (2) the validation procedure, (3) development of the text-based measure and (4) the data.

First, the reliability of the text-based measure was only assessable to a limited level due to external factors such as costly similar software and the availability of panel data. The transient error was not tested in this research as the dataset did not consist of responses collected at two points in time, which is an essential condition of testing this measurement error. Further research is needed to increase reliability.

Second, following the work of Short and colleagues (2010), there are several validation issues when using CATA. Most of them are discussed in this research except those of external validity and predictive validity. External validity is hardly possible to examine in this specific context analysis, mainly due to the limited availability of comparable research. The need to assess all the constituents of validity gives opportunities for further research. This might be a difficult task for external validity as comparable research is still limited, however, predictive validity could be analysed.

Third, it is not advisable to generate a more extensive wordlist than necessary concerning the data. In the wordlists of this study, however, some words, e.g. blah, agog and gung-ho, were not that relevant for the constructed text-based measure. Therefore it is advisable to analyse and contextualise words carefully to create even more accurate wordlists. Developing dictionaries for all job affect dimensions instead focussing on only boredom and enthusiasm will give more information about the additional value open questions can have in surveys. These emotion-specific dictionaries could give a better opportunity to analyse the discriminant

validity as it will specifically show the relationship with the related constructs for every job affect dimension. Moreover, focusing on all 20 words used in the JAWS instrument will create an overall text-based measure that, following theoretical assumptions, should have a stronger convergence with job satisfaction.

The final limitations concern the dataset used in this study, which was relatively small as it only consisted of 364 respondents. An inductive approach, which is data-driven, gives less extensive results with a small dataset. Therefore, it is also advisable to use a larger dataset for further research as it could then generate a more extensive inductively derived wordlist, which will increase the validity of the text-based measure. This could also be done by letting respondents answer with an infinite amount of words, however, the answer quality could decrease as respondents final words might be less important than their first responded words. Another limitation related to the data concerns the quality of the responses. As the answers to the semi-open question might be too diverse, it is probably better to include a question with a focus on one part of the emotion scale, for example, a question related to high activation and low arousal. Respondents can choose to answer on only one of the four parts of the emotion scale that is applicable to them. This will generate more specific adjectives. In this way, the answers will be less random and better able to be analysed on their convergence, discriminant and predictive validity. Constructing text-based measures to open questions instead of semi-open questions could also be interesting for further research. Measuring open questions will come with some difficulties as the answers will be more even more diverse and there will be many unneeded words in the responses. Complete open text files, e.g. shareholder letters (Short et al., 2010), related to job affect seems a hard task as well as this often includes formal writing styles and is not primarily based on just adjectives of job emotions. This does not mean it would not be interesting to research but the development of a valid and reliable text-based measure will be a harder task.

All in all, this study demonstrated that semi-open questions with job affect words can be used to construct a context-specific text-based measure to analyse the relationships with related job- and life-satisfaction constructs. Generating both theory-driven and data-driven dictionaries can be useful to enhance reliability and construct validity using CATA. Therefore, it is interesting to expand this study field in further CATA research by analysing the added value of open questions in similar survey instruments, which considered closed-questions as the standard. The study could function as an inspiration for scholars interested to build new research questions and hypotheses related to content analysis studies using CATA.

References

- Agho, A. O., Price, J. L., & Mueller, C. W. (1992). Discriminant validity of measures of job satisfaction, positive affectivity and negative affectivity. *Journal of occupational and organizational psychology*, 65 (3), 185-195.
- Aziri, B. (2011). Job satisfaction: A literature review. *Management Research & Practice*, 3 (4).
- Barr, P. S., Stimpert, J. L., & Huff, A. S. (1992). Cognitive change, strategic action, and organizational renewal. *Strategic management journal*, 13 (1), 15-36.
- Bateman, T. S., & Organ, D. W. (1983). Job satisfaction and the good soldier: The relationship between affect and employee "citizenship". *Academy of management Journal*, 26 (4), 587-595.
- Bench, S., & Lench, H. On the function of boredom. *Behavioral sciences*, 3 (3), 459-472.
- Borg, I., & Zuell, C. (2012). Write-in comments in employee surveys. *International Journal of Manpower*, 33 (2), 206-220.
- Bowling, N. A., & Hammond, G. D. (2008). A meta-analytic examination of the construct validity of the Michigan Organizational Assessment Questionnaire Job Satisfaction Subscale. *Journal of Vocational Behavior*, 73 (1), 63-77.
- Cable, D. M., & Judge, T. A. (1996). Person-organization fit, job choice decisions, and organizational entry. *Person-organization fit, job choice decisions, and organizational entry. Organizational behavior and human decision processes*, 67 (3), 294-311.
- Cammann, C., Fichman, M., Jenkins, D., & Klesh, J. (1979). *The Michigan organizational assessment questionnaire*. Ann Arbor, University of Michigan: Unpublished manuscript.
- Carsten, J. M., & Spector, P. W. (1987). Unemployment, job satisfaction, and employee turnover: A meta-analytic test of the Muchinsky model. *Journal of Applied Psychology*, 72, 374-381.
- CBS. (2018). *Meer psychische vermoeidheid ervaren door werk*. From Centraal Bureau voor de Statistiek: <https://www.cbs.nl/nl-nl/nieuws/2018/46/meer-psychische-vermoeidheid-ervaren-door-werk>
- Connolly, J. J., & Viswesvaran, C. (2000). The role of affectivity in job satisfaction: A meta-analysis. *Personality and individual differences*, 29 (2), 265-281.
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, 49 (1), 71-75.

- Duriau, V. J., Regeer, R. K., & Pfarrer, M. D. (2007). A content analysis of the content analysis literature in organization studies: Research themes, data sources, and methodological refinements. *Organizational research methods*, 10 (1), 5-34.
- Farrell, D. (1983). Exit, voice, loyalty, and neglect as responses to job dissatisfaction: A multidimensional scaling study. *Academy of Management Journal*, 26, 596-607.
- Gilles, I., Mayer, M., Courvoisier, N., & Peytremann-Bridevaux, I. (2017). Joint analyses of open comments and quantitative data: Added value in a job satisfaction survey of hospital professionals. *PLoS ONE*, 12(3), e0173950.
- Glerum, A., Atasoy, B., & Bierlaire, M. (2014). Using semi-open questions to integrate perceptions in choice models. *Journal of choice modelling*, 10, 11-33.
- Halbesleben, J. R., & Wheeler, A. R. (2008). The relative roles of engagement and embeddedness in predicting job performance and intention to leave. *Work & Stress*, 22 (3), 242-256.
- Heller, D., Judge, T. A., & Watson, D. (2002). The confounding role of personality and trait affectivity in the relationship between job and life satisfaction. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 23 (7), 815-835.
- Herrbach, O., & Mignonac, K. (2004). How organisational image affects employee attitudes. *Human Resource Management Journal*, 14 (4), 76-88.
- Holsti, O. R. (1969). *Content analysis for the social sciences and humanities*. Reading, MA: Addison-Wesley.
- Hom, P. W., Katerberg, R., & Hulin, C. L. (1979). A comparative examination of three approaches to the prediction of turnover. *Journal of Applied Psychology*, 64, 280-290.
- Hulin, C. L. (1968). The effects of changes in job satisfaction levels on turnover. *Journal of Applied Psychology*, 52, 122-126.
- Hulin, C. L. (1966). Job satisfaction and turnover in a female clerical population. *Journal of Applied Psychology*, 50, 280-285.
- Joffe, H., & Yardley, L. (2004). Content and thematic analysis. *Research methods for clinical and health psychology*, 56, 68.
- Judge, T. A., Weiss, H. M., Kammeyer-Mueller, J. D., & Hulin, C. L. (2017). Job attitudes, job satisfaction, and job affect: A century of continuity and of change. *Journal of Applied Psychology*, 102 (3), 356.
- Kabanoff, B., Waldersee, R., & Cohen, M. (1995). Espoused values and organizational change themes. *Academy of Management Journal*, 38 (4), 1075-1104.

- Kaplan, S. A., Warren, C. R., Barsky, A. P., & Thoresen, C. J. (2009). A note on the relationship between affect (ivity) and differing conceptualizations of job satisfaction: Some unexpected meta-analytic findings. *European Journal of Work and Organizational Psychology*, 18 (1), 29-54.
- Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. Sage publications.
- Kristof, A. L. (1996). Person-organization fit: An integrative review of its conceptualizations, measurement, and implications. 49 (1), 1-49.
- Kristof-Brown, A. L., Zimmerman, R. D., & Johnson, E. C. (2005). Consequences of individual's fit at work: A meta-analysis of person-job, person-organization, person-group, and person-supervisor fit. *Personnel psychology*, 58 (2), 281-342.
- Looker, E., Denton, M. A., & Davis, C. K. (1989). Bridging the gap: Incorporating qualitative data into quantitative analyses. *Social Science Research*, 18 (4), 313-330.
- McKenny, A. F., Short, J. C., & Newman, S. M. (2012). CAT Scanner (Version 1.0) [Software]. Available from <http://www.amckenny.com/CATScanner/>
- McKenny, A. F., Aguinis, H., Short, J. C., & Anglin, A. H. (2018). What doesn't get measured does exist: Improving the accuracy of computer-aided text analysis. *Journal of Management*, 44 (7), 2909-2933.
- Mobley, W. H. (1977). Intermediate linkages in the relationship between job satisfaction and employee turnover. *Journal of Applied Psychology*, 62 (2), 237-240.
- Mobley, W. H., Horner, S. O., & Hollingsworth, A. T. (1978). An evaluation of precursors of hospital employee turnover. *Journal of Applied Psychology*, 63, 408-414.
- Moorman, R. H. (1993). The influence of cognitive and affective based job satisfaction measures on the relationship between satisfaction and organizational citizenship behavior. *Human relations*, 46 (6), 759-776.
- Neuendorf, K. (2002). *The content analysis guidebook*. Los Angeles, CA: Sage.
- Organ, D. W. (1988). *Issues in organization and management series. Organizational citizenship behavior: The good soldier syndrome*. Lexington, MA, England: Lexington Books/D. C. Heath and Com.
- Organ, D. W., & Near, J. P. (1985). Cognition vs affect in measures of job satisfaction. *International Journal of Psychology*, 20 (2), 241-253.
- Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70, 153-163.

- Petrone, B., & Kc Whissell, C. (1988). The Dictionary of Affect in Language as a tool for the assessment of affective tone in a descriptive task. *Perceptual and Motor Skills* , 67, 789-790.
- Porter, L. W., Steers, R. M., Mowday, R. T., & Boulian, P. V. (1974). Organizational commitment, job satisfaction, and turnover among psychiatric technicians. *Journal of Applied Psychology* , 59 (5), 603-609.
- R Core Team. (2014). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <http://www.R-project.org/>
- Reijseger, G., Schaufeli, W. B., Peeters, M. C., Taris, T. W., van Beek, I., & Ouweneel, E. (2013). Watching the paint dry at work: Psychometric examination of the Dutch Boredom Scale. *Anxiety, Stress & Coping* , 26 (5), 508-525.
- Russell, J. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology* , 39, 1161-1178.
- Schaufeli, W. B., & Bakker, A. B. (2004). Job demands, job resources, and their relationship with burnout and engagement: A multi-sample study. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior* , 25 (3), 293-315.
- Schaufeli, W. B., & Salanova. (2014). Burnout, boredom and engagement at the workplace. *People at work: An introduction to contemporary work psychology* .
- Schaufeli, W., & van Rhenen, W. (2006). Over de rol van positieve en negatieve emoties bij het welbevinden van managers: Een studie met de Job-related Affective Well-being Scale (JAWS). *Gedrag & Organisatie* , 19, 323-344.
- Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods* , 1 (2), 199.
- Schmidt, F. L., & Hunter, J. E. (1999). Theory testing and measurement error.
- Schmidt, F. L., Le, H., & Ilies, R. (2003). Beyond alpha: An empirical examination of the effects of different sources of measurement error on reliability estimates for measures of individual difference constructs. *Psychological Methods* , 8, 206-224.
- Schuman, H., & Presser, S. (1996). Questions and answers in attitude surveys: experiments on question form, wording, and context. *Quantitative Studies in Social Relations* .
- Short, J. C., Broberg, J. C., Coglisier, C. C., & Brigham, K. H. (2010). Construct validation using computer-aided text analysis (CATA) an illustration using entrepreneurial orientation. *Organizational Research Methods* , 13 (2), 320-347.

- Short, J. C., McKenny, A. F., & Reid, S. W. (2018). More than words? Computer-aided text analysis in organizational behavior and psychology research. *Annual Review of Organizational Psychology and Organizational Behavior* , 5, 415-435.
- Smith, P. C. (1969). The measurement of satisfaction in work and retirement: A strategy for the study of attitudes.
- Spector, P. E. (1997). *Job satisfaction: Application, assessment, causes, and consequences* (Vol. 3). Sage.
- Spector, P. E., Bauer, J. A., & Fox, S. (2010). Measurement artifacts in the assessment of counterproductive work behavior and organizational citizenship behavior: Do we know what we think we know? *Journal of Applied Psychology* , 95 (4), 781.
- StataCorp. 2017. *Stata Statistical Software: Release 15*. College Station, TX: StataCorp LLC.
- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological bulletin*, 87(2), 245.
- Stevenson, A., & Waite, M. (2011). *Concise Oxford English Dictionary (12th ed.)*. New York, NY: Oxford University Press.
- Tett, R. P., & Meyer, J. P. (1993). Job satisfaction, organizational commitment, turnover intention, and turnover: path analyses based on meta-analytic findings. *Personnel psychology* , 46 (2), 259-293.
- Thompson, C., & Gregory, J. B. (2012). Managing millennials: A framework for improving attraction, motivation, and retention. *The Psychologist-Manager Journal* , 15 (4), 237-246.
- Thompson, E. R., & Phua, F. T. (2012). A brief index of affective job satisfaction. *Group & Organization Management* , 37 (3), 275-307.
- Tsai, C.-J. (2016). Boredom at Work and Job Monotony: An Exploratory Case Study within the Catering Sector: Boredom at work and job monotony. *Human Resource Development Quarterly* , 27, 206-236.
- Van Katwyk, P. T., Fox, S., Spector, P. E., & Kelloway, E. K. (2000). Using the Job-Related Affective Well-Being Scale (JAWS) to investigate affective responses to work stressors. *Journal of occupational health psychology* , 5 (2), 219.
- Van Saane, N., Sluiter, J. K., Verbeek, J. H., & Frings-Dresen, M. H. (2003). Reliability and validity of instruments measuring job satisfaction—a systematic review. *Occupational medicine* , 53 (3), 191-200.
- Weber, R. P. (1990). *Basic content analysis* (Vol. 49). Sage.
- Weiss, D. J., Dawis, R. V., & England, G. W. (1967). Manual for the Minnesota satisfaction questionnaire. *Minnesota studies in vocational rehabilitation*.

- Weiss, H. M. (2002). Deconstructing job satisfaction: Separating evaluations, beliefs and affective experiences. *Human resource management review* , 12 (2), 173-194.
- Weiss, H. M., & Cropanzano, R. (1996). Affective events theory: A theoretical discussion of the structure, causes and consequences of affective experiences at work. *Research in Organizational Behavior* , 19, 1-74.
- Wolfe, R. A., Gephart, R. P., & Johnson, T. E. (1993). Computer-facilitated qualitative data analysis: Potential contributions to management research. *Journal of Management* , 19 (3), 637-660.
- Wood, A. M. (1998). *The effects of teacher enthusiasm on student motivation, selective attention, and text memory*. London, Ontario, Canada: University of Western Ontario.
- Yaniv, E., Lavi, O. S., & Siti, G. (2010). Person-organisation fit and its impact on organisational citizenship behaviour as related to social performance. *Journal of General Management* , 36 (2), 81-89.
- Zehner, F., Sälzer, C., & Goldhammer, F. (2016). Automatic coding of short text responses via clustering in educational assessment. *Educational and psychological measurement* , 76 (2), 280-303.