CHARLES UNIVERSITY IN PRAGUE

FACULTY OF HUMANITIES

BACHELOR'S THESIS

Liberal Arts and Humanities

An Investigation of Effort-Reward Imbalance and Cognitive Inhibition in the Context of Higher Education

by Jackson David Ellison

Supervisor: Timothy Wells, MSc., Ph.D.

Prague 2019

DECLARATION

I hereby declare that no portion of the work referred to in this thesis has been submitted in support of an application of another degree, or qualification thereof, or for any other university or institute of learning.

I declare that this thesis is my own independent work. All the used material and literature has been duly referenced and quoted.

In Prague, 26 June, 2019 *signature:* ____________________
Jackson David Ellison

Acknowledgements

I would like to thank my supervisor for his guidance, encouragement and support.

I would like to thank Veronika, without whom I would not have been able to continue.

I would also like to thank my colleagues, for comradeship along the way.

Table of Contents

KEY WORDS

effort-reward imbalance, cognitive inhibition, Stroop task, general health, higher education

1 Introduction

Stress is associated with a wide variety of psychological and somatic pathologies, ranging from depression and problematic drinking behavior (Bobák, 2005), to cardiovascular disease (Cohen, Janicki-Deverts, & Miller, 2007), obesity (Kivimäki et al., 2006), HIV progression (Patterson et al., 2013; Kemeny & Schedlowski, 2007), stroke (Tsutsumi, Kayaba, Kario & Ishikawa, 2009) and cancer (Kemeny & Schedlowski, 2007). Accordingly, stress and its causes demand attention as significant threats to public health. Whereas much is known about stress and its causes in the workplace, less is known about the causes and consequences of stress in the academic environment.

1.1 The Whitehall Studies

Whereas the relationship between stress and somatic pathology has been acknowledged for some time (Kemeny & Schedlowski, 2007), it is only since the Whitehall studies that stress has become a major focus of epidemiological research (Palmer, 2015). The Whitehall studies were two longitudinal studies on the social determinants of health. The first began in the 1960s and ended ten years later. The second Whitehall study began in the 1980s, and is ongoing (Dich, Rozing, Kivimäki, & Doan, 2019). The purpose of both studies was to track the development of somatic pathology across time. A wide variety of data was collected from participants about their income, diet, workplace, personal relationships, smoking habits, physical health, height, childhood, etc. The cohorts of both studies consisted exclusively of British Civil Service employees, with each study including more than 10,000 participants (Marmot et al., 1991). The central finding of both studies is that even when all other relevant differences are controlled for, lower social rank consistently predicts increased mortality (Marmot & Sapolsky, 2014). In the British Civil Service, all employees have access to public healthcare of the same quality. Additionally, every employee has a rank which determines their work and level of control within the organization. As such, the British Civil Service serves as an ideal situation to analyze, because healthcare access is uniform across the sample, while the level of control is variable (Marmot & Sapolsky, 2014).

1.2 Uncontrollability

There are multiple explanations for the finding that low rank is associated with higher mortality, and they are not mutually exclusive. One point that they converge on is the role of uncontrollability in the relationship between workplace and health outcomes. A meta-analysis of over two-hundred studies concerning short-term human stress response found that uncontrollability is one of the two causes of stress in humans (Dickerson & Kemeny, 2004). Because the ethics of eliciting stress in human populations over the long-term is prohibitory, experimental manipulations of human stress response over the long-term are lacking (Shields, Sazma, & Yonelinas, 2016; Rugulies, Aust, & Madsen, 2016). Still, many prospective cohort studies have examined human stress response over the long-term, and they converge on the same conclusion (Kivimäki & Siegrist, 2016; Siegrist, 2017).

In short-term and long-terms studies of human stress response, uncontrollability is defined as the inability to control the outcome of a situation, especially when the outcome of the situation determines the achievement of a salient goal (Dickerson & Kemeny, 2004). For example, white noise does not cause stress, by itself. However, when white noise is affecting participants' concentration, where participants need to concentrate to achieve some goal, the white noise acts as a stressor (Dickerson & Kemeny, 2004). Alternatively, controllable noise has been found to be less stressful than noise which is uncontrollable (Glass & Singer, 1972).

1.3 Effort-Reward Imbalance

There are two models which have been applied to the Whitehall studies and similar studies, to explain their results. These models are not mutually exclusive, and they are widely understood to be complimentary (Bosma, Peter, Siegrist, & Marmot, 1998; Ostry, Kelly, Demers, Mustard, & Hertzman, 2003; Kivimäki & Siegrist, 2016). Both models measure uncontrollability in the workplace. One of these two models is the demand-control model. It concerns the demands on a worker in their workplace, and the amount of control that they have over their work. As such, the demands-control model is very much restricted to the workplace, the dynamics of the workplace and a person's working situation. Although the demand-control model has been applied in other contexts, it is less frequently applied to other contexts than the effort-reward imbalance model.

The other model used to explain workplace stress is the effort-reward imbalance model. The effort-reward imbalance model assesses the amount of effort a worker expends in their employment, and the amount of reward that they attain as a result (Siegrist, 2017). There is a third component of this model which has not been as widely validated, termed overcommitment (van Vegchel, de Jonge, Bosma, & Schaufeli, 2005). Unlike overcommitment, high effort and low reward have been consistently correlated with increased mortality across European populations (Dragano et al., 2017). Particularly, the ratio between effort and reward has shown a strong effect (van Vegchel et al., 2005). With reference to uncontrollability, this makes sense. If one is rewarded for the effort that one exerts, one can be said to have control over the situation. Whereas, if one's efforts do not result in progress towards a salient goal, i.e. if one's high efforts do not result in the attainment of a proportionally high reward, one clearly lacks control. High effort reward imbalance, i.e. high effort accompanied by low reward, indicates high uncontrollability. Accordingly, high effort-reward imbalance is associated with stress and the symptoms of stress.

The effort-reward imbalance model was first proposed in the 1980s, and in the intervening decades has been correlated with a wide variety of pathologies in multiple prospective cohort studies (van Vegchel et al., 2005). Most of this research has been conducted in Western Europe, although researchers such as Dr. Bobák have investigated effort-reward imbalance in the Czech Republic and Russia, and other post-socialist countries (Pikhart et al., 2001; Bobák at al., 2005; Chen et al., 2016). Recently, research involving the effort-reward imbalance model has been conducted in Asia and South America, replicating the results of research conducted in Western Europe (Li, 2016; Tsutsumi, 2016; Ortiz & Juárez-García, 2016).

Across studies, high effort-reward imbalance is associated with overall increased mortality. In working populations in the Czech Republic, effort-reward imbalance has been associated with low self-rated health (Pikhart et al., 2001), depression and problematic drinking behavior (Bobák et al., 2005). Recently, a study of high school students in China linked effort-reward imbalance in school to worse self-rated health and suicidal ideation (Li, Shang, Wang, & Siegrist, 2010; Shang, Li, Li, Wang, & Siegrist, 2013). Whereas the effort-reward imbalance model has been primarily used in studies of working populations, calls have been made for the model to be adapted to other situations (Siegrist, 2017). Aside from the adaption of the model to the school setting in China, the model has been used to study stay-at-home mothers in Germany. It was found that stay-at-home mothers, like working people, suffer worse health where there is higher effort-reward imbalance, indicating the broad applicability of the construct (Sperlich, Arnhold-Kerri, Siegrist, & Geyer, 2012).

1.4 Causality

1.4.1 Does Depression Cause Stress?

A key finding in the research concerning uncontrollability, which is also frequently listed as a limitation in studies of effort-reward imbalance, is the dependence of uncontrollability upon participants' perceptions of uncontrollability (Koolhaas et al., 2011). Where a situation is actually uncontrollable, lower perceptions of uncontrollability are associated with a lower stress response. Whereas, in the same situation, higher perceptions of uncontrollability are associated with a higher stress response (Kemeny, 2003; Kemeny & Gruenewald, 2000). Given the strong association between effort-reward imbalance and depressive disorders, this is a salient point. It may be that individuals with depressive disorders tend to perceive situations as uncontrollable, and therefore become stressed as a result of their perceptions (Gabrys, Tabri, Anisman, & Matheson, 2018).

1.4.2 Does Stress Cause Depression?

Although it likely is the case that people with major depressive disorder suffer stress as a result of their perceptions, it is nevertheless likely that stress causes depression. A wealth of literature supports the latter assertion, literature which provides a comprehensive neuroendocrinological explanation of the relationship between uncontrollable stress and depressive symptomatology. Briefly stated, stress is proinflammatory. Short-term inflammation can prevent an infection, but persistent inflammation over a longer period can result in a range of different diseases. Chronic stress can cause chronic inflammation, inflammation over a longer period of time. Chronic inflammation activates an adaptive response, termed sickness behavior.

Sickness behavior is characterized by anhedonia and altered cognition, among other symptoms. Some researchers have drawn a parallel between the symptomatology of sickness behavior and depressive symptomatology, inferring that this mechanism may account for the relationship between depression and stress (Bucci, Marques, Oh, & Harris, 2016; Kemeny, 2003; Cohen et al., 2007). Some researchers have also speculated that this mechanism might mediate the relationship between effort-reward imbalance and depression (Bellingrath & Kudielka, 2016).

1.4.2.1 Does Effort-Reward Imbalance Cause Depression?

Additionally, a recent review of studies on the relationship between effort-reward imbalance and depression found that effort-reward imbalance was consistently associated with the onset of depression across multiple prospective studies (Rugulies et al., 2016). What is meant by prospective study in this case, is a study in which data is taken from participants from the onset, and then again at a later time, to assess the development of the dependent variable, which is usually a negative health outcome. In the case of the aforementioned prospective studies, effort-reward imbalance was the independent variable and depression was the dependent variable. In most of the studies reviewed, the time from the initial measurement to the follow-up ranged from one to five years (Rugulies et al., 2016). While causal inferences cannot necessarily be drawn from prospective studies such as the aforementioned, multiple prospective studies which demonstrate an association between effort-reward imbalance and the development of depressive disorders do point in the direction of a causal relationship. This is especially true in light of mechanistic explanations, of which there are many (Bellingrath & Kudielka, 2016).

1.5 Executive Function

In addition to depression, neuroscientists have found that even mild experiences of acute uncontrollable stress can result in a loss of executive function (Arnsten, 2009). Where this stress is chronic, impairments of executive function over the long-term are observed (Bucci et al., 2016; Arnsten, 2009).

Executive functions are those functions that allow consideration, focus, emotional control, and goal-directed action, broadly speaking (Shields et al., 2016). Different categorical schemes exist, but most authors consider executive function in terms of three core functions. These three core functions are working memory, inhibition, and cognitive control (Diamond, 2013). Support for the idea that there are three core executive functions primarily comes from factor analysis and neuroimaging studies. It has also been found that damage to specific brain regions is associated with deficits in specific executive functions (Shields et al., 2016). Deficits in executive functioning are associated with a wide variety of pathologies, such as addiction, schizophrenia, and depression. Worse executive functioning is associated with poorer academic performance and unemployment. Remarkably, executive functioning is also a better predictor of school readiness than IQ (Diamond, 2013).

1.5.1 Inhibition

Inhibition, one of the three core executive functions, is associated with a variety of disorders in children and adolescents. These disorders include conduct disorder, obsessive compulsive disorder, attention deficit hyperactivity disorder and various learning disabilities (Dempster & Corkill, 1999). Inhibition is subdivided by different authors in a variety of different ways (Joormann, Yoon, & Zetsche, 2007). Often, inhibition is divided into two components, response inhibition and cognitive inhibition (Shields et al., 2016). Explanations of these two components vary somewhat, and overlap with components of different categorical schemes. Nevertheless, the components are most commonly described as follows.

Response inhibition is the ability to inhibit prepotent motor responses, i.e. to inhibit behavior. Cognitive inhibition is ability to selectively attend to or ignore information. For cognitive inhibition, this includes the ability to ignore unwanted thoughts or memories. Definitions of cognitive inhibition are somewhat variable, and overlap with definitions of selective inhibition, a third component of inhibition in other categorical schemes (Diamond, 2013). Some authors differentiate between behavioral inhibition and response inhibition (Nigg, 2016), while other authors use the terms behavioral inhibition and response inhibition interchangeably (Grillon et al., 2016).

Although a variety of perspectives can be found throughout the literature concerning inhibition and its components (Hallion, Tolin, Assaf, Goethe, & Diefenbach,, 2017; Diamond, 2013), factor analysis has revealed response inhibition and cognitive inhibition to be indistinguishable, at least in samples of healthy young adults (Shields et al., 2016). Moreover, neuroimaging studies indicate that inhibitory control of action and attention are on the neural level similar if not indistinguishable (Diamond, 2013). In short, although inhibition is a hot topic in cognitive psychology, and a great diversity of speculative views exist, inhibition is actually a well understood and thoroughly studied cognitive function, which has been integrated into the explanations of a variety of psychopathologies (MacLeod, 2007; Johnson, 2007).

1.5.1.1 Cognitive Inhibition and Rumination

Going forward, only cognitive inhibition will be discussed, not because cognitive inhibition is in any discernible way distinct from response inhibition, but rather only because cognitive inhibition is the term used to refer to inhibitory control where thoughts or emotions are concerned. For this reason, because it pertains to thoughts and emotions, cognitive inhibition features heavily in theoretical discussions of affective disorders, such as major depressive disorder and generalized anxiety disorder. Deficits in cognitive inhibition have been associated with both (Richard-Devantoy et al., 2012; Gohier et al., 2009; Hallion et al., 2017).

Multiple explanations exist for the associations between cognitive inhibition and various psychopathologies. One explanation is that cognitive inhibition facilitates emotional control, so that in its absence emotional control is no longer possible (Joormann & Gotlib, 2010). The concept of rumination plays a vital role in this explanation; it is thought that rumination, which is a symptom of depression, reflects the inability to inhibit negative thoughts and memories. In the absence of attentional or otherwise inhibitory control, patients cannot effectively regulate their emotions. Accordingly, worse cognitive inhibition is associated with difficulty disengaging from repetitive negative thoughts (Joormann et al., 2007) and "the inability to disengage from negative emotional information" (Gabrys et al., 2018).

Deficient cognitive inhibition is positively associated with rumination and negatively associated with cognitive reappraisal (Joormann & Gotlib, 2010). Rumination is the tendency to attend to the same thoughts over and over again. Rumination is a pattern of thinking, characterized by the continued recycling of identical thoughts. It is maladaptive, in that it leads to the exaggeration of negative emotions and experiences. Rumination can prolong depression (Joormann & Gotlib, 2010), and precipitate the onset of depression (Joormann & Gotlib, 2010).

Cognitive reappraisal, on the other hand, refers to actively switching from one perspective to another. Cognitive reappraisal is thought to be a healthy form of emotional regulation, because it allows situations to be reinterpreted such that they elicit a less negative or even positive affective state (Joormann & Gotlib, 2010).

There is a negative relationship between cognitive reappraisal and depression, whereas there is a positive relationship between rumination and depression. Deficient cognitive inhibition is associated with less frequent use of cognitive reappraisal as a form of emotional regulation, whereas deficient cognitive inhibition is associated with more frequent use of rumination as a form of emotional regulation. Recalling that "rumination in response to negative events and negative mood states increases the risk for the onset of a depressive episode" (Joormann & Gotlib, 2010), and given that rumination is associated with cognitive inhibition, it may be that decreases in cognitive inhibition could precipitate a depressive episode.

According to a recent meta-analysis, comprising 47 studies which employed the Stroop task as a measure of cognitive inhibition, the association between cognitive inhibition and depression is characterized by a large effect. Moreover, the effect increases with the severity of depression, such that the greater the depression severity the lower the cognitive inhibition (Epp, Dobson, Dozois & Frewen, 2012). Here it should be noted that the relationship between cognitive inhibition and rumination is at present only correlational, although it is theoretically substantiated. The claim that there is a causal relationship between cognitive inhibition and rumination may therefore be unjustified. There could be a third factor (Joormann & Gotlib, 2010).

1.5.1.2 Cognitive Inhibition and Stress

Whereas there is some relationship between uncontrollable stress and executive function, and although cognitive inhibition is a component of one of the three core executive functions, the nature of the relationship between stress and cognitive inhibition is not clear. This is primarily due to a lack of research on the subject. It is also partially due to definitional issues, viz. the various ways in which cognitive inhibition is defined with respect to the other components.

At least two articles have been published on the relationship between cognitive inhibition and stress. One of those is a meta-analysis which employs a categorical scheme which is too ambiguous to be interpreted with reference to the broader literature on cognitive inhibition (Shields et al., 2016). The other is a small-scale study that was conducted on undergraduates in Israel, to look at the relationship between cognitive inhibition, depression, anxiety and stress. Surprisingly, they did not find a significant relationship between cognitive inhibition and depression, cognitive inhibition and anxiety, or cognitive inhibition and stress (Ajilchi & Nejati, 2017). This is surprising for two reasons. The first is that the association between cognitive inhibition and depression has shown a large effect, in a number of other studies, with reference to the same general measure of cognitive inhibition that the authors used (Epp et al., 2012). Some authors have also found an association between cognitive inhibition and generalized anxiety disorder, via the same measure (Hallion et al., 2017). One explanation for the nonsignificance of the results in question, is that their study was non-clinical, viz. all of their participants were undergraduates in Tehran, Iran (Ajilchi & Nejati, 2017). While it is possible that there were some students who had been diagnosed with clinical depression, in their sample of 488 participants, there is no indication that any of them were. Where the relationship between cognitive inhibition and depression is analyzed, the largest effect is found in clinical studies, studies in which some of the participants have been diagnosed with a depressive disorder (Epp et al., 2012). It may be that, in the nonclinical sample of undergraduates, the effect was too small to be visible.

1.6 Aims

Based on the epidemiological and psychological literature, it is clear that there is a relationship between uncontrollability and depression. It is also clear that uncontrollability causes reductions in executive function. Moreover, multiple meta-analyses have consistently associated depression with deficits in executive function (Epp et al., 2012). Based on the finding that depression is associated with deficits in executive functioning, some researchers have speculated that reduced executive function can cause depression. The question that arises from this mess of associations and theoretical postulations is whether or not reduced executive function therefore mediates the relationship between uncontrollability and depressive disorders.

The present study addresses this question in the form of a much more specific hypothesis. The form that the question takes in the present study, is whether effort-reward imbalance causes depression by reducing cognitive inhibition. In other words, is cognitive inhibition one of the mechanisms that mediates the causal relationship between effort-reward imbalance and depression? Admittedly, making causal claims is difficult, especially on the basis of correlations. Regardless, cross-sectional studies such as this one, which aim to assess the relationship between a pathology and its probable causes, are often necessary to initiate and orient research in the direction of more granular accounts. In this respect, the present research project is not without precedent. Correlations between effort-reward imbalance and depression have paved the way for psychoneuroendocrinological explanations of the etiology of depression (Bellingrath & Kudielka, 2016). What the present study hopes to contribute, is a hypothesis that accounts for one of the mechanisms that mediates the relationship between uncontrollability and depressive disorders. Namely, the present study aims to describe a cognitive mechanism that, like the psychoneuroendocrinological mechanisms, explains the causal relationship between effort-reward imbalance and depression. Additionally, the present study aims to reproduce the results of similar studies which have investigated the relationship between effort-reward imbalance and psychopathology across diverse populations.

The aims of this study are therefore twofold. First, this study aims to assess the relationship between effort-reward imbalance and cognitive inhibition, to support or reject the hypothesis that chronic uncontrollability reduces cognitive inhibition. Second, this study aims to reproduce the findings which theoretically underpin the first hypothesis. Namely, the present study aims to reproduce the finding that effort-reward imbalance and general control are associated with self-rated health, depression, anxiety, and other measures. What differentiates the present study from other studies which have employed the effort-reward imbalance model, is that the present study concerns the Czech student population. For the most part, the effort-reward imbalance model has been applied to working populations. Only somewhat recently has its application been extended to other populations, such as stay-at-home mothers and high school students (Siegrist, 2017).

The present study aims to accomplish the first aim, i.e. assessing the relationship between uncontrollability and cognitive inhibition, by measuring effort-reward imbalance and cognitive inhibition, and subsequently running a statistical test to assess whether there is a relationship. Effort-reward imbalance is usually assessed via a questionnaire. Accordingly, in this study, effort-reward imbalance will be assessed via an effort-reward imbalance questionnaire. Cognitive inhibition is most commonly assessed via the Stroop task, and it will be assessed via the Stroop task in the present study. The second aim of this study, which is to reproduce the correlations that other researchers have found between uncontrollability and psychopathology, will be achieved by administering measures of uncontrollability and psychopathology. Namely, questionnaires to measure effort-reward imbalance, control, self-rated health, depression, anxiety, and problematic drinking behavior will be administered. Statistical tests will be subsequently executed to assess the correlations.

2 Methods

2.1 Introduction

This study consisted in the administration of a few psychometrics and a cognitive task. These measures will be explained in turn, before delving into the minutiae of the procedures.

2.2 Questionnaires

2.2.1 Effort-Reward Imbalance Questionnaire

A questionnaire to assess effort-reward imbalance was administered (Siegrist, 1996). The effort-reward imbalance questionnaire is standardized, and around half of all studies which measure effort-reward imbalance employ the standardized questionnaire (Van Vegchel, 2005). However, the standardized effort-reward imbalance questionnaire was designed to be administered to workers. To be administered to other populations, it needs to be adapted.

One recent study used one such adapted effort-reward imbalance questionnaire as a measure in a population of Chinese high school students. The researchers reproduced the same associations that have been found in working populations. Namely, among the Chinese high school students, effort-reward imbalance was associated with worse self-rated health and increased suicidal ideation (Li et al., 2010; Shang et al., 2013). Because the present study also concerns students, the questionnaire from the Chinese studies was adapted to the Czech undergraduate population. The English version of the questionnaire that was administered to high school students in that study can be found in its appendix (Li et al., 2010). For the present study, the questionnaire from the Chinese study was slightly rephrased, so that it could be administered to university rather than high school students. After the questionnaire was adapted, it was translated into Czech.

The effort-reward imbalance questionnaire that was administered in this study consisted of four questions in the effort component and eleven questions in the reward component. The overcommitment component of the questionnaire was not included, for theoretical and practical reasons. The theoretical reason is that the overcommitment component of the effort-reward imbalance questionnaire has undergone multiple theoretical reimaginings since the questionnaire was introduced, such that what the overcommitment component actually measures has changed over time (Van Vegchel et al., 2005). This is especially problematic given that many researchers do not use the standard effort-reward imbalance questionnaire, even when administering it to working populations (Van Vegchel, 2005). Rather than using the standard questionnaire, some researchers use other measures as proxies for the different components of the effort-reward imbalance construct (Kivimäki et al., 2006). Accordingly, insofar as these proxies are based on the theoretical definitions of the components, given that the theoretical definition of the overcommitment component has changed over time, results concerning the overcommitment component across studies are not necessarily comparable. The practical reason that the overcommitment component of the questionnaire was not included was because it would add extra questions to the overall count. Due to the component's theoretical issues, its addition to a list of questions which was already long could not be justified. To reiterate, the effort-reward imbalance model consists of three components: overcommitment, effort and reward. The effort component and the reward component were administered, but the overcommitment component was not. On the effort-reward imbalance questionnaire, all questions are answered on a 5-point Likert scale.

During analysis, each participants' score on the effort component was divided by their score on the reward component, resulting in their effort-reward ratio. This effort-reward ratio is a measure of effort-reward imbalance, and effort-reward imbalance is a better predictor of pathology than either the effort or the reward score taken individually. In other words, neither effort nor reward are as highly correlated with psychopathology as effort-reward imbalance is (Sperlich et al., 2012; Li et al., 2010). This may be because high effort-reward imbalance represents high uncontrollability, i.e. it represents a situation in which effort is exerted towards a goal where the rewards are largely unaffected by the effort exerted. Whereas, neither high effort nor low reward necessarily imply uncontrollability by themselves.

2.2.2 Control Questionnaire

Another questionnaire that was administered to measure uncontrollability was a questionnaire to assess perceived control (Bobák, Pikhart, Hertzman, Rose, & Marmot, 1998).

Unlike the effort-reward imbalance questionnaire, the control questionnaire has not been widely administered. Rather, the control questionnaire was only used once, in a study of the Russian population in the 1990s, during the economic transition. What that study found was largely in accordance with the claims already outlined. Lower control was correlated with worse self-rated health. It should be noted that low self-rated health is correlated with depression and other psychopathologies. Moreover, and as was mentioned in that study, self-rated health has been found to predict health outcomes above and beyond physicians' assessments.

The control questionnaire is made up of two components, which in this study will be treated as two separate measures. One component assesses general control, and the other component assesses control over health. The components will be treated as two separate measures, because there is not a solid theoretical relationship between the two components. An English language version of the questionnaire was taken from the appendix of the study in which it was first used. The questionnaire was translated from English into Czech. Because of the general nature of the questionnaire, it did not need to be altered in any way before being administered to university students. The questionnaire consists of nine questions overall. There are six questions in the general control component and three questions in the component which measures perceived control over their health. Both components of the control questionnaire are scored on a 6-point Likert scale.

2.2.3 General Health Questionnaire

A questionnaire to assess self-rated health was also administered. As has already been mentioned, low self-rated health is consistently associated with high effort-reward imbalance. This is a finding of studies which investigated working populations, stay-at-home mothers, and most recently, Chinese high school students (Siegrist, 2008; Sperlich et al., 2012; Li et al., 2010). Therefore, a reproduction of this finding among Czech university students would constitute a useful contribution to the literature, as well as reproduce the findings of other researchers in similar settings. As was already mentioned, low self-rated health has also been associated with low control, and low control over health, in Russia in the 1990s (Bobák et al., 1998).

There are multiple valid and widespread measures of self-rated health. In one study "self-rated health was assessed by a single item, 'How would you say your health is?,' using 4-point Likert scale ratings of excellent, good, fair, and poor" (Li et al., 2010). Some studies have used the SF-36 (Bobák et al., 1998), similar studies have also used other measures (Gådin & Hammarström, 2000). The SF-36 consists of 36 questions and therefore may be excessively lengthy. A shorter questionnaire which is no less valid is the 12-Item General Health Questionnaire (GHQ-12) (Kashyap & Singh, 2017; Goldberg, & Blackwell, 1970). As the name suggests, it consists of 12 questions. There are multiple possible scoring methods. However, in the present study the questionnaire was scored on a 4-point Likert scale. It can also be scored such that the best two or worst two responses to a question are scored with a 1 or a 0, respectively. Instead, the Likert scale scoring method was chosen, because the Likert scale scoring method is more normally distributed and better represents variations in severity (Del Pilar Sánchez-López & Dresch, 2008).

Unlike the effort-reward imbalance questionnaire and the general control questionnaire, the GHQ-12 did not need to be translated into Czech for this study. A translation of the questionnaire into Czech was already available.

2.2.4 Problematic Drinking Behavior Questionnaire

Given the number of studies that have been carried out in the Czech Republic on the relationship between workplace stress and problematic drinking behavior (Bobák, 2005), a measure of problematic drinking behavior was also administered. The Short Michigan Alcohol Screening Test (SMAST) did not need to be translated, because it was already available in Czech. The SMAST consists of 13 yes or no questions (Selzer, Vinokur, & van Rooijen, 1975). Each positive answer adds one point to the participant's score, while negative answers do not add any points to the score. The higher the SMAST score, the higher the likelihood of alcoholism. A score between zero and two indicates that there is no problem. A score of three indicates a borderline alcohol problem, and a score of four or more indicates the presence of problematic drinking behavior. One study found that "at the SMAST's suggested cutoff score of 3, the average sensitivity was .68 and specificity was .74" (Minnich et al., 2019).

2.2.5 Depression Questionnaire

A questionnaire to measure depression was administered. It was tempting to use the self-report Hamilton Depression Inventory, due to its use elsewhere in the literature (Kasch et al., 2002), but its use has been criticized (Bagby et al., 2004). One alternative is the self-rating version of the Montgomery–Åsberg Depression Rating Scale (MADRS-S), which has been suggested as an alternative (Bagby et al., 2004; Heo, Murphy, & Meyers, 2007); however, its correlation with physicians' assessments is a bit shaky (Cunningham et al., 2011). Because it is highly intercorrelated with the MADRS-S (Svanborg & Åsberg, 2001), the Beck Depression Inventory (BDI) is another option, but its twenty plus questions were too many. In light of these considerations, the Patient Health Questionnaire (PHQ-9) emerged as ideal. It has been recommended as a "diagnostic tool in the research of population-based samples where face-to-face diagnostic interviews are not available," and consists of only nine questions (Martin et al., 2006, p. 76). It is scored on a 4-point Likert scale. When an interview with a healthcare professional is used as the point of reference, "a PHQ-9 score ≥10 [has] a sensitivity of 88% and a specificity of 88% for major depression" (Kroenke, Spitzer, & Williams, 2001, p. 606).

Additionally, a longitudinal study found that "each 1-point increase on the PHQ-9 depression severity measure led to a 6% increased risk of mortality," and that the questionnaire "showed a similar power to detect the association of depressive symptoms and long-term mortality [as] clinical diagnoses" (Martin-Subero et al., 2017, p. 280). A version of the PHQ-9 was already available in Czech, and so did not need to be translated ("Zdravotní Dotazník," n.d.).

2.2.6 Anxiety Questionnaire

Generalized anxiety disorder and depression are often comorbid (Robert & Hirschfeld, 2001), and generalized anxiety disorder has been associated with cognitive inhibition (Hallion et al., 2017). Accordingly, a measure of anxiety may serve to reinforce measurements of depression. On the other hand, the relationship between anxiety and cognitive inhibition is not altogether straightforward (Grillon et al., 2016). As such, the measurement of anxiety and its relationship to cognitive inhibition may serve to clarify relationships that have been observed, in addition to potentially clarifying other relationships in the present study.

There are a few measures of generalized anxiety disorder throughout the literature. The Hamilton Anxiety Rating Scale has been used in other studies concerning the relationship between GAD and cognitive inhibition (Hallion et al., 2017). However, the Hamilton Anxiety Rating Scale is not available in Czech. To be administered to Czech students it would need to be translated. The self-administered version of the Hamilton Anxiety Rating Scale is also no longer available online

The Generalized Anxiety Disorder 7-item scale (GAD-7) is highly correlated with the Hamilton Anxiety Rating scale. Additionally, there is "agreement between the classification groups generated by both scales," when patients are put into groups based on anxiety severity (Garcia-Campayo et al., 2010; Ruiz et al., 2011). The GAD-7 is therefore a suitable substitute for the Hamilton Anxiety Rating scale. The GAD-7 is also the shorter of the two scales, and therefore preferable, given the number of questions that were already set to be administered. A Czech version of the GAD-7 was available, and therefore a translation did not need to be made for the present study (Marková, 2018).

The GAD-7 is scored on a 4-point Likert scale. The author of the measure has suggested that scores of 5, 10, and 15 represent mild, moderate, and severe levels of anxiety, respectively. According to some researchers, "a score of 10 or greater on the GAD-7 represents a reasonable cut point for identifying cases of [generalized anxiety disorder]" (Spitzer, Kroenke, Williams, & Löwe, 2006). Using a threshold score of 10 or higher, the GAD-7 has a sensitivity of .89 and a specificity of .82 (Kroenke, Spitzer, Williams, Monahan, & Löwe, 2007). At a cut-off score of 8 for generalized anxiety disorder, the GAD-7 has a sensitivity of .83 and a specificity of .84 (Plummer, Manea, Trepel, & McMillan, 2016). In the present study, the GAD-7 will be used to assess the severity of anxiety, rather than its presence or absence.

2.2.7 Demographic Questionnaire

A demographic questionnaire was also administered. Participants were asked their age, sex, nationality, and university of attendance. Included in the demographic questionnaire were questions pertaining to other variables which might affect participants' performance on the cognitive task that will be described below.

2.3 Cognitive Task

2.3.1 Introduction

Stroop first published a description of the Stroop task in 1935, and in the intervening years the Stroop effect has become "one of the best known phenomena in all of cognitive science and indeed in psychology more broadly" (MacLeod, 2015). Accordingly, there is a massive body of literature on the Stroop task and its correlates. Although Stroop clearly outlined his methods, subsequent administrations of the Stroop task have been accompanied by innumerable variations on the original (Homack, 2004).

The original Stroop task consisted of two components. There was the color reading component, and the interference component. Both components were administered on sheets of paper. For the color component, participants were given a paper with different colored boxes on it. The boxes could be one of five colors: blue, red, green, brown, or purple. There were one hundred boxes in total. Participants were instructed to read out the colors of the boxes as quickly as possible. Researchers recorded the amount of time participants took to name the colors of all of the boxes, the number of correct responses, and the number of incorrect responses.

The second component of the original Stroop task was termed the interference component. This component was identical to the color component, except for the content of the paper sheet that participants were instructed to read from. Rather than one hundred color boxes, participants were given a sheet with one hundred words, in different colors. The words were "blue," "red," "green," "brown," and "purple." Each word was colored in one of the other four colors. For example. "the word 'blue' was printed in red, green, brown, and purple inks" (Stroop, 1935, p. 648). Like the color component, participants were instructed to read out the colors of the words on the sheet. Participants were not instructed to read the words themselves, but rather the color that the words were printed in. As in the first component, the amount of time participants took to finish reading the items, the number of correct responses and the number of incorrect responses were recorded.

The amount of time taken to complete the first component was compared to the amount of time taken to complete the second component. What was found was that participants could read out color boxes faster than they could read colored words, when the words and their colors were incongruous. Stroop called this phenomenon the interference effect, also known as the Stroop effect. It was thought that word reading was an automatic response, such that naming an incongruously colored word would require the inhibition of the automatic response to read out the word (Flaudias & Llorca, 2014). For example, if a participant was presented with the word "red," colored green, the participant would have to inhibit their automatic response to read "red," so that they could read out the word's color. Accordingly, the Stroop effect measures inhibition. The greater the difference between a participant's scores on the color and interference component, the lower their inhibition. Conversely, the lower the difference between a participant's scores on those two components, the higher their inhibition.

2.3.2 Variations

Although the method of scoring the cognitive task was clearly described in the 1935 paper, a great variety of new scoring methods have proliferated since that time. The import of this is that the way in which the Stroop task is scored has not been standardized. Different authors use different methods of scoring the Stroop task, and scoring methods can affect purported results. Administration of the task itself has also undergone substantial changes. At present, the task is not standardized. Nevertheless, there are widespread tendencies in the task's administration, and there have been clear general changes over time.

"In the typical clinical version of the Stroop test, participants complete a pure block of neutral trials (reading the color-words and/or naming the ink-color of neutral words) followed by a pure block of incongruent trials." (Bélanger, Belleville, & Gauthier, 2010, p. 582). The exact configurations of the different blocks vary from author to author. For example, whereas Stroop originally used solid squares of color for the neutral block (Stroop, 1935), some researchers use lines of the same letter, or asterisks (Siegrist, 1997; Battisti et al., 2010). The size of the font in the interference condition is also variable across studies (Bélanger et al., 2010; Siegrist, 1997).

A major shift in the administration of the Stroop task has come with computers. In the original Stroop task, participants read lists of words or colors and vocalized the names as quickly as possible. In this version, experimenters would listen, time and write down the vocalizations. Now, the Stroop task is usually administered on a computer (Homack, 2004). In computerized versions, the colors or words are shown one at a time. Rather than responding aloud, participants press a key on a keyboard. In this version, the computer records the participants' response times and errors. Computerized Stroop tasks have been shown to elicit interference effects comparable to the traditional Stroop task. However, the relationship between the traditional and computerized Stroop task is not altogether clear. Nevertheless, use of the computerized Stroop task is widespread (Penner et al., 2012).

In 1965, one author claimed to have found "eleven formulas for deriving scores from the three basic Stroop scores … in the Stroop literature" (Jensen, 1965, p. 401). Since that time even more methods of scoring the Stroop task have been created. There was not a standard scoring formula for the Stroop task in 1965, and there is not a standardized method for scoring the Stroop task now (Jensen, 1965; Homack, 2004). However, some scoring methods are more popular than others (Scarpina & Tagini, 2017). Similarly, the computerized Stroop task is not standardized, although some recommendations are more popular than others (Homack, 2004).

To quote a recent review of the Stroop task literature, "we recommend that a scoring method for the [Stroop task] should fulfill two main requirements. First, both accuracy and speed must be computed for all [experimental] conditions. And secondly, a global index must be calculated to relate the performance in the incongruous condition to reading words and color naming abilities. The first requirement can be achieved by counting the number of correct answers in each condition in within a fixed time (Amato et al., 2006; Valgimigli et al., 2010; Brugnolo et al., 2015). The second requirement can be achieved by subtracting the [word] score and [color] score from [the interference] score" (Scarpina & Tagini, 2017). They also noted that none of the studies that they reviewed fulfilled both of those requirements.

2.3.3 Design

According to recommendations from the above review as well as popular convention, the version of the Stroop task that was administered in the present study consisted of three components. There was a color component, a word component, and an interference component. In addition to these three, there was also a practice component which preceded the other components. The cognitive task was created in PsychoPy (Peirce et al., 2019).

Before the start of each component, an instructional screen appeared. There were different instructions for each of the three main components. Instructions, and all other text, were in Czech. To continue from the instruction screen to the task, participants needed to press a key confirming that they had finished reading the instructions. Before the start of each component, participants were reminded of the potential responses. There were four potential responses to every stimulus. The keys corresponding to these responses stayed the same in every component. The key "a" mapped to the color red, the key "s" mapped to the color blue, the key "j" mapped to the color yellow and the key "k" mapped to the color green.

Each of the three main components lasted for 45 seconds. Each component consisted in the sequential presentation of stimuli. For example, in the color component, one color would appear on the screen, e.g. green. The participant would respond by pressing the key corresponding to the color green, which is the "k" key. Following the participant's response, the stimulus would disappear, interspersed by a blank screen lasting for 0.5 seconds. Following the blank screen, another stimulus would appear, etc. After 45 seconds of this, the component would end, and the instructional screen for the next component would appear. The only difference between the three main components were their instructions and stimuli. The font for all stimuli was Arial, with a font size of approximately 24.

The color component's stimuli were lines of consisting of the letter X repeated five times, XXXXX. The line would appear in one of the four colors: red, blue, yellow or green. Participants were instructed to press the key corresponding to the color on the screen. This component is a computerized version of the color component in the original Stroop color task, except that in this version a string of colored characters rather than a colored box was presented. There is little to indicate that the difference matters, but the use of colored character strings rather than colored boxes is common among the recent administrations of the Stroop task.

The word component's stimuli were words. One of the four color words would be presented, and the font color of every word was black. Participants were instructed to press the key corresponding to the word on the screen. For example, if the Czech translation of the word "blue" appeared onscreen, the correct response would be to press the "s" key, which corresponds to the color blue. To reiterate, all the words would be colored black. Participants would read the word and press the key corresponding to the word that they read.

The interference component's stimuli were like the stimuli administered in the interference component of the original Stroop task (Stroop, 1935). One of the four color words would appear, but always colored incongruously. For example, the Czech translation of the word "blue" might appear in the colors red, yellow or green. The word blue would not appear in the color blue. Participants were instructed to press the key corresponding to the color of the word, not the word itself. For example, if the Czech translation of the word "red" appeared on the screen, and its font color was green, the correct response would be to press the key corresponding to the color green.

To reiterate, all three of the main components were 45 seconds. This setup is keeping with the recommendations of a recent review and in contradistinction to Stroop's original formulation. In the original Stroop task, there was a sheet with a certain number of words or colors on it, and participants would read the words as quickly as they could. As such, participants' scores would be based on the amount of time that they took to say the words or names of the colors. Contrary to this way of calculating participants scores, in the version of the Stroop task that was administered in the present study, participants' scores were calculated with reference to the number of correct responses that they made. The purpose of this, according to the review, is to account for both accuracy and speed. By considering the number of correct answers that participants make in a 45 second period, both speed and accuracy are accounted for, in one measurement (Scarpina & Tagini, 2017). Whereas, were Stroop's original method retained, speed and accuracy could not be so seamlessly expressed in a single score. Rather, further calculations would have to be made to account for speed and accuracy, to bring them together into a single score. Even though summing correct answers does not account for incorrect answers directly, because every response incurs a blank screen that persists for 0.5 seconds, incorrect answers are indirectly accounted for when scores are summed over a 45 second period.

In addition to the three main components, there was also a practice block. The purpose of the practice block was to assure that participants knew which key on the keyboard was associated with which color, and to make sure that participants understood the task.

The practice component was almost identical to the color component, except that it was not programmed to terminate after 45 seconds. The instructional screen was the same, the keys were the same, and the stimuli presented were the same. The difference between the practice component and the other components was that the practice component was not timed. The practice component would only terminate after the participant made five correct answers in a row. Every time the participant made an incorrect answer, the count would restart. To prevent the situation in which a participant continued to get incorrect answers and got stuck in an infinite loop, a maximum number of stimuli that could be presented during the practice component was set. After some number of responses, if the participant still had not made five correct answers in a row, the practice component would terminate automatically.

To recap, the Stroop task was administered on a computer. The task consisted of three main components, in addition to a practice component to assure that participants understood the nature of the task and knew which keys mapped to which colors. The three main components were the color component, the word component, and the interference component. Every participant completed the components in the same order. Before each component, an instructional screen would appear, and participants were required to press a key to confirm that they had read the instructions before continuing. The three main components were set to last 45 seconds each, in accordance with a recent review's recommendations (Scarpina & Tagini, 2017). The scores for each component were calculated by summing the number of correct responses that the participant made over the course of 45 seconds.

2.3.4 Scoring

The score on the Stroop task can be calculated in a wide variety of ways. One author found eleven different methods of scoring the Stroop task, and that was in 1965. Since then, even more methods have entered the literature. As such, choosing one method over any other presents problems, not least of which is the problem of generalizability and comparison. Even if the most theoretically comprehensive scoring method was chosen, there is no guarantee that the method would measure the same phenomenon as other scoring methods. Different scoring methods may measure different phenomena (Scarpina & Tagini, 2017). Generalizing about cognitive inhibition from an uncommon scoring method could therefore be in error, since the

Stroop effect has been and will likely continue to be defined by other scoring methods. Moreover, the investigation of cognitive inhibition necessarily involves some experimental conventions. To disregard the conventions by which earlier conclusions have been come upon jeopardizes the project's grounding in the literature. For these reasons, the most common method of scoring the Stroop task was been chosen. This happens to be Golden's method (Scarpina & Tagini, 2017).

Golden's scoring method is calculated based on the scores of all three components. Somewhat more complicated than addition and subtraction, the formula will not be outlined here. Before analysis, Goldman's scoring formula was programmed into R. This function took the three scores from the three components as inputs and returned a single Stroop score as output. This score was taken to be a measure of cognitive inhibition. Higher scores represent higher cognitive inhibition. Lower scores represent lower cognitive inhibition. Low cognitive inhibition is associated with a variety of psychopathologies (Johnson, 2007).

2.4 Study Participants

Participants were recruited at the Jinonice campus of the Faculty of Humanities, Charles University. All participants were either graduate or undergraduate students at Charles University. Of the 39 participants, 2 were Slovak and 37 were Czech. The information sheet, consent form, cognitive task, and questionnaires were written in Czech.

2.5 Data Processing

In addition to participants' Stroop scores, all questionnaire data was imported into R. Because questionnaire data was obtained on paper, the questionnaire data was manually input into a computer for analysis. The questionnaire data would first be input into a spreadsheet. Formulas were inserted into the spreadsheets to calculate total scores on the questionnaires, and automatically perform computations such as reverse scoring, to mitigate potential operator error. The spreadsheets were two pages each. The second page would present the data that had been input into the first page, in summarized and simplified form. This second page would thereafter be exported as a .csv and subsequently imported into R for analysis. This process prevented potential errors and streamlined the process, from data acquisition to analysis.

2.6 Ethics

For the questionnaires and the cognitive task, each participant was assigned a unique, anonymizing identification code. Participants' answers to the questionnaires as well as participants' scores on the cognitive task were fully anonymized. During analysis, participants were only analyzed with reference to their identification codes. Prior to data collection, each participant was provided with an information sheet and a consent form. According to the information sheet, participants were informed that their questionnaire answers and their scores on the cognitive task were anonymous. Only upon completion of the consent form could participants start the cognitive task.

The present study was conducted in accordance with the ethical guidelines of the American Psychological Association and the British Psychological Society (American Psychological Association, 2017; The British Psychological Society, 2014).

2.7 Statistical Tests

Three statistical tests were performed, to test the hypothesis and its different aspects.

First, a t-test was run to determine whether effort-reward imbalance affects cognitive inhibition. Participants with scores 1 standard deviation above or below the median Stroop score were excluded from this analysis, to exclude the scores of participants who were distracted or misunderstood the task. This was keeping with the recommendation that scores below 1 standard deviation on any of the components are suspect (Golden, Espe-Pfeifer, & Wachsler-Felder, 2000). Among the remaining participants, a median split was performed, which divided the participants into two groups based on their effort-reward imbalance ratios. This is similar to other studies, which have divided participants into two groups according to their effort-reward ratios (Siegrist, 2008; Sperlich et al., 2012). There were two groups, one above the median effort-reward imbalance ratio and one below the median effort-reward imbalance ratio. A t-test was then performed, comparing the scores on the Stroop task of these two groups. This was a test of the main hypothesis, which is that effort-reward imbalance is associated with cognitive inhibition.

The second statistical procedure performed was a correlation matrix. No participants were excluded from the correlation matrix. The correlation matrix included eleven variables: problematic drinking behavior, effort, reward, effort-reward imbalance, depression, self-rated health, general control, control over health, anxiety, age and sex. The purpose of the correlation table was to examine the relationships between all of the variables involved, to assess correspondence with the broader literature.

The third statistical test performed was a multiple regression. There were eight predictor variables: problematic drinking behavior, depression, self-rated health, general control, control over health, anxiety, age and sex. Three outliers were excluded from this analysis. The dependent variable was effort-reward imbalance. The purpose of this test was to clarify the relationship between effort-reward imbalance and the other variables that were measured, to clarify the relationship assessed with the t-test.

All data analysis was performed in R (R Core Team, 2018).

3 Results

3.1 Questionnaires

The scores on the Short Michigan Alcoholism Screen Test are a sum of affirmative answers. Every affirmative answer equates to 1 point. There are 13 questions in all, M = 3.15, SD = 1.09, range = 1 - 6.

The scores on the effort-reward imbalance questionnaire are measured on a 5-point Likert scale. Some of the questions are reverse scored. There are 4 questions in the effort component, M = 8.62, SD = 3.07, range = 3 - 16. There are 11 questions in the reward component, M = 29.74, SD = 4.25, range = 23 - 37. Effort-reward imbalance is the ratio of effort and reward, where effort is the numerator and reward is the denominator, M = 0.30, SD = 0.14, range = 0.10 - 0.70.

The Patient Health Questionnaire is a measure of depression, scored on a 4-point Likert scale. There are 9 questions in all, M = 7.33, SD = 4.31, range = 0 - 18.

The 12-item General Health Questionnaire is a measure of self-rated health. It is scored on a 4-point Likert scale, M = 13.13, SD = 6.22, range = 6 - 30.

The control questionnaire consists of two components, which in the present study are treated as two separate measures, both scored on a 6-point Likert scale. The general control component is made up of 6 questions, M = 18.49, SD = 3.62, range = 12 - 28. The control over health component consists of 3 questions, M = 10.97, SD = 2.27, range = 5 - 15.

The General Anxiety Disorder questionnaire consists of 7 questions and is scored on a 4-point Likert scale, M = 5.41, SD = 4.95, range = 0 - 19.

Participants provided their ages, M = 21.56, SD = 2.43, range = 19 - 27. There were 11 males and 27 females. The sex of 1 participant is unknown.

3.2 Cognitive Task

The Stroop task is divided into three components. Scores on each component are the sum of the participant’s correct responses within a 45 second period. On the color component, participants press the key corresponding to the color of a character string, M = 26.21, SD = 6.59, range = 10 - 37. On the word component, participants press the key corresponding to the word on the screen, M = 25.54, SD = 6.27, range = 13 - 38. On the interference component, participants press the key corresponding to the color of the word on the screen, M = 18.36, SD = 9.69, range = 0 - 33.

Overall scores on the Stroop task were computed with Golden's formula (as cited in Scarpina & Tagini, 2017), M = 5.64, SD = 8.58, range = -15.48 – 17.60.

3.3 T-test

Participants (n = 39) with a score on the Stroop task more or less than one standard deviation away from the median were excluded. The remaining participants (n = 28) were split into two groups, above and below the median effort-reward ratio. Similar divisions have been made in other studies (Van Vegchel et al., 2005; Sperlich et al., 2012)

A t-test was run to examine the difference in mean Stroop scores between the group of participants above the median effort-reward ratio (M = 9.64, SE = 3.37) and the group below the median effort-reward ratio (M = 7.95, SE = 3.23). Surprisingly, the mean score on the Stroop task was higher in the group with a higher effort-reward ratio than it was in the group with the lower effort-reward ratio, although this difference was not significant, $t(25.95) = -1.35$, $p = .19$, with an effect of $r = .26$.

3.4 Correlations

The correlation matrix (n = 39) includes eleven variables. In Table 1, "Alcohol" refers to the Short Michigan Alcoholism Screening Test (SMAST), which is a measure of problematic drinking behavior. "Effort" and "Reward" refer to the effort and reward components of the effort-reward imbalance questionnaire, respectively. "ERI" stands for effort-reward imbalance, which is the product obtained from dividing a participant's score on the effort component by their score on the reward component of the effort reward imbalance questionnaire. "PQH-9" stands for Patient Health Questionnaire, which is the questionnaire that was chosen to measure depression in this study. "GHQ-12" refers to the 12-item General Health Questionnaire, which is a measure of self-rated health. "GControl" refers to the general control component of the control questionnaire, and "HControl" refers to the control over health component of the same questionnaire. "GAD-7" stands for the Generalized Anxiety Disorder 7 questionnaire, which is a measure of anxiety. Because some of the variables' distributions were non-normal, the Spearman rank correlation coefficient rather than the Pearson correlation coefficient was used.

Table 1

Spearman Correlations

	1	2	3	4	5	6	7	8	9	10
1. SMAST										
2. Effort	-.10									
3. Reward	.21	-.54*								
4. ERI	-.16	.93*	-.78*							
5. PHQ-9	.14	.48*	-.38*	.48*						
6. GHQ-12	.03	.51*	-.41*	.52*	.78*					
7. GControl	.08	-.39*	.50*	-.41*	-.58*	-.48*				
8. HControl	-.07	-.17	.46*	-.33*	-.21	-.25	.22			
9. GAD-7	.14	.33*	-.26	.33*	.73*	.86*	-.48*	-.31		
10. Age	.16	.03	.28	-.07	-.01	-.02	.23	.18	-.03	
11. Sex	-.30	.19	-.46*	.34*	.23	.21	-.15	-.30	.21	.00
Mean	3.15	8.62	29.74	0.30	7.33	13.13	18.49	10.97	5.41	21.56
SD	1.09	3.07	4.25	0.14	4.31	6.22	3.62	2.27	4.95	2.43

*p<.05

3.5 Multiple Regression

Multiple regression was applied to participant data (n = 35) to assess whether scores on the other measures significantly predicted effort-reward imbalance. The predictors included scores on the Short Michigan Alcoholism Screening Test (SMAST), the 9-item Patient Health Questionnaire (PHQ-9), the General Health Questionnaire (GHQ-12), the general control component of the control questionnaire, the control over health component of the control questionnaire, the 7-item Generalized Anxiety Disorder (GAD-7) questionnaire, age and sex.

In Table 2, "GControl" refers to the general control component of the control questionnaire, and "HControl" refers to the component which assesses control over health. The

PHQ-9 is a measure of depression, while the GHQ-12 is a measure of patients' self-rated health.

The model was found to be a significant predictor of effort-reward imbalance, at $F(8,26) = 2.43$, $p = .041$, $r^2 = 0.4284$. The model's adjusted r^2 was .2526. In other words, the model significantly predicts 42% of the variance in effort-reward imbalance, while none of the predictors show themselves to be individually significant.

Table 2

Multiple Regression

	B	SE B	β	t	P
(Constant)	0.153302	0.228188		0.672	0.5076
SMAST	-0.007620	0.022076	-0.06982741	-0.345	0.7327
PHQ-9	0.003117	0.007553	0.11370844	0.413	0.6832
GHQ-12	0.010874	0.005652	0.54106847	1.924	0.0654
GControl	-0.002736	0.006804	-0.08040415	-0.402	0.6909
HControl	-0.010934	0.009652	-0.20815010	-1.133	0.2677
GAD-7	-0.011793	0.007387	-0.44869300	-1.597	0.1225
Age	0.008618	0.008799	0.16465361	0.979	0.3364
Sex	0.071724	0.049232	0.28240937	1.457	0.1571

4 Discussion

The t-test was nonsignificant ($p = .19$) and therefore failed to support the hypothesis that effort-reward imbalance is associated with cognitive inhibition. There are multiple possible explanations for this outcome. Naturally, one of them is that the hypothesis is simply incorrect. However, there are a few reasons to believe that an effect may have been present nevertheless.

Across studies, the association between depression and scores on the Stroop task is somewhat inconsistent (Ajilchi & Nejati, 2017). Some researchers have attributed the inconsistency between different experiments which employ the Stroop task to methodological issues. Specifically, the variety of scoring methods and variations in the design of the Stroop task itself have been identified as problems. Recalling the variation in scoring and administration of the Stroop task that was outlined in the Methods section, it has been noted that, "although mean effect sizes reveal broad impairment on neuropsychological measures of [executive function] for patients with [major depressive disorder], there is also high variability in effect sizes across studies. There may be multiple sources of this variability. … [For example,] the standard neuropsychological version of the Stroop task, with separate blocks of neutral and incongruent stimuli, is easier than versions in which trial types are intermixed … . In addition, studies often report different dependent measures for the same task," i.e. different researchers provide scores which are the outcomes of different scoring methods (Snyder, 2013, p. 101). Were the Stroop paradigm and scoring method of this study different, it stands to reason that the results would have been different as well. However, the methodology of the present study was sound, and based on convention. The point is nevertheless worth keeping in mind for future research.

Another point is that many studies investigating the relationship between depression and cognitive inhibition involved clinical participants. A lower proportion of the studies which did not include clinical participants found an effect between cognitive inhibition and depression (Epp et al., 2012). This study may have failed to find an effect for a similar reason. In other words, studies on the relationship between depression and cognitive inhibition which included individuals on both sides of the spectrum, i.e. healthy and severely depressed, more often found an effect between depression and cognitive inhibition. Likewise, it may be that the present study lacked participants on the more severe end of the spectrum of effort-reward imbalance. More specifically, it is not difficult to imagine reasons why students suffering from severe effort-reward imbalance, and all of the somatic and psychological pathologies which often accompany it, might be less likely to be present on campus. The present study's sample consisted entirely of students who were present on a university campus at the time that the study was conducted. Only students who happened to be on the campus on those days were recruited. Accordingly, it is possible that sampling error led to a situation in which university students with high effort-reward imbalance were proportionally underrepresented. Although this would not necessarily call into question the study's results, it would obscure the effect of a relationship between effort-reward imbalance and cognitive inhibition, in much the same way that a sample of participants who are not severely depressed might lead to a failure to demonstrate an association between depression and cognitive inhibition.

This potential error is not unique to the present study, but rather plagues this kind of research in general. In the epidemiological literature, a sampling error resulting from such a cause is termed the healthy worker effect. The healthy worker effect refers to the finding that mortality is lower among workers than among the general population, supposedly because the sick are less likely to join the workforce, and because people who are unhealthy are more likely to leave the workforce. Resultantly, comparisons of the workforce and the general population can be misleading (Li & Sung, 1999). This is not a problem in the present study, because there is only one population that was investigated, and it was not compared to the general population. However, the premises that the healthy worker effect is based on are relevant here. If an aspect of this study's hypothesis is correct, which is that effort-reward imbalance is a phenomenon with consequences in student populations comparable to those consequences in working populations, it would follow that students with high effort-reward imbalance tend to be physically and psychologically unwell. Accordingly, it could be reasonably expected that such students would be less likely to attend class, and more likely to drop out of university. For this reason, sampling at a university campus might tend exclude such students, i.e. students who are experiencing or have experienced high effort-reward imbalance.

In much the same way and for much the same reason as the effect of the association between depression and cognitive inhibition is more easily observed among severely depressed patients, the relationship between effort-reward imbalance and cognitive inhibition might be more easily observed among students who experience worse effort-reward imbalance. Such students were less likely to have taken part in this study than their healthier counterparts, if the hypothesis that effort-reward imbalance is associated with psychological and somatic pathologies in the Czech student population is to be believed.

The correlation matrix did show correlations between effort-reward imbalance and nearly every measure of somatic and psychological pathology. Effort-reward imbalance was most highly correlated with negative self-rated health ($r_s = .52$). Effort-reward imbalance was also correlated with depression ($r_s = .48$), and negatively correlated with general control ($r_s = -.41$) and control over health ($r_s = -.33$). The correlation with control over health is less notable, and because control over health is not theoretically significant in relation to effort-reward imbalance, it neither confirms nor calls into question any aspect of the central hypothesis. On the other hand, the correlations of effort-reward imbalance with self-rated health, depression and general control are noteworthy. The negative correlation with general control further suggests the relevance of uncontrollability to the explanation of effort-reward imbalance, although weakly.

Notably, the correlations of effort-reward imbalance with depression and negative self-rated health reproduce the findings of other studies conducted in similar populations. A relationship between effort-reward imbalance and negative self-rated health has been found among Chinese high school students (Li et al., 2010). Additionally, depression is one of the pathologies most consistently associated with effort-reward imbalance, across diverse populations (Siegrist, 2008). In the Czech Republic, a relationship has been found between effort-reward imbalance and depression, in working populations (Pikhart et al., 2004). The findings of the present study therefore reinforce and contribute to the findings of other researchers that have studied the relationships between effort-reward imbalance, depression and self-rated health.

However, the negative correlation ($r_s = -.16$) between effort-reward imbalance and problematic drinking behavior might be at odds with previous research, research that has been conducted in the Czech Republic and elsewhere (Bobák, 2005). While the effect is sufficiently small that it can be considered inconsequential, the lack of a positive correlation between effort-reward imbalance and problematic drinking behavior is surprising, in light of previous research. There are a few potential reasons for this.

One of them is that problematic drinking behavior, as measured by questionnaires, is symptomatic of different psychological states in university students and working people. In other words, it is possible that drinking patterns which are often symptomatic of psychological pathology in working populations are not symptomatic of any psychological pathology among university students. There are two explanations which support this interpretation. One is that students are in a different situation than workers, with different norms, such that drinking behaviors which result in problems for workers do not entail the same adverse social consequences for students. The second explanation for the disparity is that the average age of the student population is much lower than the average age of the working populations which have been studied. What is termed problematic drinking behavior may have different causes and consequences for younger and older people, such that drinking behavior that is problematic among older populations neither indicates pathology nor similarly leads to it among younger populations (O'Malley, 2004; Blow & Barry, 2002).

The second main reason for this finding may have to do with the nature of the measure itself. The maximum score on the measure is 12, and the minimum score is 0, however more than half of the participants scored a 3 on the questionnaire. According to the questionnaire's recommendations, a score of 3 or more indicates a borderline alcohol problem, i.e. the majority of the sample have a borderline alcohol problem, according to the questionnaire's guidelines. If we assume that most Czech university students do not have a borderline alcohol problem, what this indicates is that the measure may not be optimized to represent the nuances of problematic drinking behavior among Czech university students. This could be because the measure is not suited to the university environment or because the measure is not suitable to be administered in the Czech Republic, or due to both of these reasons. In any case, other Czech researchers have noted the complexity of problematic drinking behavior in the Czech Republic and the other post-socialist countries, advocating for more in-depth and comprehensive measurement (Bobák et al., 2015).

Sex was found to be correlated with effort-reward imbalance. To be clear, the table shows that being a woman is correlated with higher effort-reward imbalance ($r_s = .34$). There are findings to suggest that uncontrollability at home results in heart problems for women, but not for men. Conversely, for men, it has been shown that stress at work is a greater predictor of heart problems than stress at home (Kivimäki et al., 2006). This is probably due to the fact that a higher proportion of women than men, in those studies, were responsible for managing the home. In line with this point, researchers have noted that some studies fail to take into account that men and women often perform different jobs. This is one of the reasons that sex differences as far as effort-reward imbalance is concerned are not altogether clear. It is also worth noting that “the societal value of female participation in the labor market is often undervalued. Even if women are doing the same job as men, their work is more frequently judged as banal and less important” (Starke & Niedhammer, 2002, p. 69). The Academy is not necessarily excepted from this trend.

In the present study, being a woman was also found to be negatively correlated with reward ($r_s = -.46$). This effect was larger than the correlation with effort-reward imbalance. This finding contrasts the results of a recent study of Chinese physicians, which found that women rated their rewards more highly than men (Li, Yang, & Cho, 2006). However, a meta-analysis of studies involving effort-reward imbalance in European populations found that “reward scores do not differ according to gender in a consistent way” (Siegrist et al., 2004, p. 1489). In light of this disparity across different populations, it may be that the relationship between sex and reward differs between populations. The results of the present study neither firmly support nor contradict this claim. In short, the association between sex and reward is ambiguous, and the results of the present study have not reduced this ambiguity. The correlation may have been simply random, especially due to the proportionally small number of males ($n = 11$) in the sample ($n = 39$).

On the other hand, general control was negatively associated with every symptom of stress that was measured. General control was negatively associated with depression ($r_s = -.58$), negative self-rated health ($r_s = -.48$), and anxiety ($r_s = -.48$). In other words, uncontrollability was found to be positively correlated with depression, negative self-rated health, and anxiety. These associations corroborate the broader research, where uncontrollability is considered a primary cause of stress (Koolhass et al., 2011). Stress, in turn, is considered a potential cause of depression, negative self-rated health and anxiety (Sapolsky, 2015). It is interesting that general control was not notably correlated with problematic drinking behavior ($r_s = .08$). This may be for the same reasons that effort-reward imbalance was not highly correlated with problematic drinking behavior, which were already elaborated.

Last but not least, there was a negative correlation between effort-reward imbalance and general control ($r_s = -.41$). If effort-reward imbalance and general control are both considered to be measures of uncontrollability, then such a relatively small correlation might be surprising.

However, it should be kept in mind that uncontrollability is considered to be one of the main causes of stress in humans. Some researchers even consider it to be one of the two main causes of stress; in one meta-analysis, social-evaluative threat is listed as the other main cause (Dickerson & Kemeny, 2004). In another article, unpredictability is posited to be the other main cause of stress. In the latter paper, the two main causes of stress, unpredictability and uncontrollability, are even considered to be overlapping criteria, such that uncontrollability is not only one of the two main causes of stress but also inseparable from the other cause as well (Koolhaas et al., 2011). Taking all of this into account, and recognizing the breadth of uncontrollability as a construct, it therefore may not be surprising that two measures of uncontrollability may not be highly correlated, especially when one of those measures only concerns a particular kind of uncontrollability. As such, it is possible to reconcile the finding of a relatively low correlation with the belief that both general control and effort-reward imbalance are measures of uncontrollability, when it is recognized that the latter measure only concerns a subset of a much larger series of phenomena.

However, it should also be mentioned that some researchers do not look at effort-reward imbalance in terms of uncontrollability, but rather in terms of other theoretical frameworks. For example, the original author of the effort-reward imbalance model considers the stress that results from effort-reward imbalance to be caused by the violation of "an 'evolutionary old' grammar" (Siegrist, 2009, p. 307). He considers this evolutionary grammar to be the rule of reciprocity, positing that the violation of this rule is what results in stress. In other words, he believes that there is an expectation hardwired into every person, that if they expend effort for someone, that their efforts will be rewarded. When employees, parents or students are not rewarded for their efforts, this expectation is violated. This explanation is unsatisfactory for two reasons. The first reason is that it is purely speculative. The second reason is that it is not grounded in the broader literature on stress, which has consistently and empirically demonstrated the relationship between endocrine markers of stress and other variables, e.g. cortisol and uncontrollability (Dickerson & Kemeny, 2004; Sapolsky, 2012).

For these reasons, the relatively small correlation between general control and effort-reward imbalance neither supports nor calls into question the theoretical perspective which posits effort-reward imbalance as an instance of uncontrollability.

The third statistical procedure performed was a multiple regression. The independent variables were problematic drinking behavior, depression, self-rated health, general control, control over health, anxiety, age and sex. The dependent variable was effort-reward imbalance. The model was a significant predictor of effort-reward imbalance, at $p = .041$, with an effect of $r^2 = .428$. However, none of the model's variables were found to be significant predictors of effort-reward imbalance individually. This is likely due to the number of predictor variables included in the model. Still, self-rated health was close to passing the threshold, at $p = .065$.

5 Conclusion

A relationship was not found between cognitive inhibition and effort-reward imbalance, however there are reasons to be believe that the effect might have been obscured. The sampling procedure could have precluded the observation of an effect, if there was one. Future research should be informed by the methodological considerations that were outlined in the Discussion.

With respect to the correlations between the psychometric measures, this study constitutes a valuable contribution to the literature. First, it demonstrates an association between effort-reward imbalance and depression in a university setting, reproducing findings from workplace studies. Second, it represents the successful adaption of the effort-reward imbalance questionnaire to the university student population. In both of these respects, this study has evidenced Siegrist's suggestion that the deleterious consequences of effort-reward imbalance can be observed across diverse populations, inside and outside of the workforce (Siegrist, 2017).

In addition to this, the present study has also reproduced the findings of the study of the Russian population in the late-1990s (Bobák et al., 1998), in accordance with the broader theoretical literature, by demonstrating an association between uncontrollability and the symptoms of chronic stress. Namely, the present study has reproduced the findings that uncontrollability is associated with depression and worse self-rated health.

With respect to sex, age and drinking behavior, results were inconclusive. Previous studies have not noted a relationship between effort-reward imbalance and age, for the most part (Li et al., 2010), such that the lack of a relationship between age and effort-reward imbalance is unsurprising. Some studies have reported that there is a relationship between sex and effort-reward imbalance, although this relationship is likely mediated by other factors, and is at present ambiguous. One somewhat surprising finding of the present study is a lack of an association between effort-reward imbalance and problematic drinking behavior, and the absence of a strong association between depression and problematic drinking behavior. Previous studies in the Czech Republic have found an association between effort-reward imbalance and problematic drinking behavior, and problematic drinking behavior and depressive disorders (Bobák, 2005; Bobák et al., 2006). While sampling may have played a role in this finding, there is also the likelihood that the norms which govern the drinking behavior of university students are different than those that govern the drinking behavior of working people. As such, the same indicators may indicate different things in these two different populations.

In all, this study failed to demonstrate its central hypothesis, but at the same time succeeded in reproducing the results of other researchers. Even the failure to prove the central hypothesis reproduced other researchers' results (Ajilchi & Nejati, 2017). Two novel measures were introduced, a new effort-reward imbalance questionnaire and a Czech translation of the control questionnaire. Both measures were associated with the expected variables, reproducing the results of similar studies (Li et al., 2010; Shang et al., 2013; Bobák et al., 1998).

6 References

Ajilchi, B., & Nejati, V. (2017). Research Paper: Executive Functions in Students With Depression, Anxiety, and Stress Symptoms. *Basic and Clinical Neuroscience Journal, 8*(3), 223–232. doi:10.18869/nirp.bcn.8.3.223

American Psychological Association (2017). *Ethical Principles of Psychologists and Code of Conduct*. Retrieved Jun. 16, 2019, from https://www.apa.org/ethics/code/index

Arnsten, A. F. T. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. *Nature Reviews Neuroscience, 10*(6), 410–422. doi:10.1038/nrn2648

Bagby, R. M., Ryder, A. G., Schuller, D. R., & Marshall, M. B. (2004). The Hamilton Depression Rating Scale: Has the Gold Standard Become a Lead Weight? *American Journal of Psychiatry, 161*(12), 2163–2177. doi:10.1176/appi.ajp.161.12.2163

Battisti, R. A., Roodenrys, S., Johnstone, S. J., Pesa, N., Hermens, D. F., & Solowij, N. (2010). Chronic cannabis users show altered neurophysiological functioning on Stroop task conflict resolution. *Psychopharmacology, 212*(4), 613–624. doi:10.1007/s00213-010-1988-3

Bellingrath, S., & Kudielka, B. M. (2016). Psychobiological Pathways from Work Stress to Reduced Health: Naturalistic and Experimental Studies on the ERI Model. In *Work Stress and Health in a Globalized Economy* (pp. 145–170). doi:10.1007/978-3-319-32937-6_7

Blow, F. C., & Barry, K. L. (2002). Use and misuse of alcohol among older women. *Alcohol Research and Health, 26*(4), 308-315.

Bobák, M., Pikhart, H., Hertzman, C., Rose, R., & Marmot, M. (1998). Socioeconomic factors, perceived control and self-reported health in Russia. A cross-sectional survey. *Social Science & Medicine, 47*(2), 269–279. doi:10.1016/s0277-9536(98)00095-1

Bobák, M. (2005). The association between psychosocial characteristics at work and problem drinking: a cross-sectional study of men in three Eastern European urban populations. *Occupational and Environmental Medicine, 62*(8), 546–550. doi:10.1136/oem.2004.017202

Bobák, M., Pikhart, H., Pajak, A., Kubinova, R., Malyutina, S., Sebakova, H., … Marmot, M. (2006). Depressive symptoms in urban population samples in Russia, Poland and the Czech Republic. *British Journal of Psychiatry, 188*(04), 359–365. doi:10.1192/bjp.188.4.359

Bobák, M., Malyutina, S., Horvat, P., Pajak, A., Tamosiunas, A., Kubinova, R., … Marmot, M. G. (2015). Alcohol, drinking pattern and all-cause, cardiovascular and alcohol-related mortality in Eastern Europe. *European Journal of Epidemiology, 31*(1), 21–30. doi:10.1007/s10654-015-0092-8

Bosma, H., Peter, R., Siegrist, J., & Marmot, M. (1998). Two alternative job stress models and the risk of coronary heart disease. *American Journal of Public Health, 88*(1), 68–74. doi:10.2105/ajph.88.1.68

Bucci, M., Marques, S. S., Oh, D., & Harris, N. B. (2016). Toxic Stress in Children and Adolescents. *Advances in Pediatrics, 63*(1), 403–428. doi:10.1016/j.yapd.2016.04.002

Bélanger, S., Belleville, S., & Gauthier, S. (2010). Inhibition impairments in Alzheimer's disease, mild cognitive impairment and healthy aging: Effect of congruency proportion in a Stroop task. *Neuropsychologia, 48*(2), 581–590. doi:10.1016/j.neuropsychologia.2009.10.021

Chen, S.-W., Peasey, A., Stefler, D., Malyutina, S., Pajak, A., Kubinova, R., … Pikhart, H. (2016). Effort–reward imbalance at work, over-commitment personality and diet quality in Central and Eastern European populations. *British Journal of Nutrition, 115*(07), 1254–1264. doi:10.1017/s0007114515005516

Cohen, S., Janicki-Deverts, D., & Miller, G. E. (2007). Psychological Stress and Disease. *JAMA, 298*(14), 1685. doi:10.1001/jama.298.14.1685

Cunningham, J. L., Wernroth, L., von Knorring, L., Berglund, L., & Ekselius, L. (2011). Agreement between physicians' and patients' ratings on the Montgomery–Åsberg Depression Rating Scale. *Journal of Affective Disorders, 135*(1-3), 148–153. doi:10.1016/j.jad.2011.07.005

Del Pilar Sánchez-López, M. & Dresch, V. (2008). The 12-Item General Health Questionnaire (GHQ-12): Reliability, external validity and factor structure in the Spanish population. *Psicothema 20*(4), 839-843.

Dempster, F. N., & Corkill, A. J. (1999). Interference and Inhibition in Cognition and Behavior: Unifying Themes for Educational Psychology. *Educational Psychology Review, 11*(1), 1–88. doi:10.1023/a:1021992632168

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology, 64*(1), 135–168. doi:10.1146/annurev-psych-113011-143750

Dich, N., Rozing M.P., Kivimäki, M. & Doan, S.N. (2019). Life Events, Emotions, and Immune Function: Evidence from Whitehall II Cohort Study. *Behavioral Medicine*. doi: 10.1080/08964289.2019.1570072

Dickerson, S. S., & Kemeny, M. E. (2004). Acute Stressors and Cortisol Responses: A Theoretical Integration and Synthesis of Laboratory Research. *Psychological Bulletin, 130*(3), 355–391. doi:10.1037/0033-2909.130.3.355

Dragano, N., Siegrist, J., Nyberg, S. T., Lunau, T., Fransson, E. I., Alfredsson, L., … Kivimäki, M. (2017). Effort–Reward Imbalance at Work and Incident Coronary Heart Disease. *Epidemiology, 28*(4), 619–626. doi:10.1097/ede.0000000000000666

Epp, A. M., Dobson, K. S., Dozois, D. J. A., & Frewen, P. A. (2012). A systematic meta-analysis of the Stroop task in depression. *Clinical Psychology Review, 32*(4), 316–328. doi:10.1016/j.cpr.2012.02.005

Flaudias, V. and Llorca, P.-M., 2014. A brief review of three manipulations of the Stroop task focusing on the automaticity of semantic access. *Psychologica Belgica, 54*(2), 199–221. doi: 10.5334/pb.am

Gabrys, R. L., Tabri, N., Anisman, H., & Matheson, K. (2018). Cognitive Control and Flexibility in the Context of Stress and Depressive Symptoms: The Cognitive Control and Flexibility Questionnaire. *Frontiers in Psychology, 9*. doi:10.3389/fpsyg.2018.02219

Garcia-Campayo, J., Zamorano, E., Ruiz, M. A., Pardo, A., Perez-Paramo, M., Lopez-Gomez, V., … Rejas, J. (2010). Cultural adaptation into Spanish of the generalized anxiety disorder-7 (GAD-7) scale as a screening tool. *Health and Quality of Life Outcomes, 8*(1), 8. doi:10.1186/1477-7525-8-8

Glass, D., C., & Singer, J., E. (1972). Behavioral Aftereffects of Unpredictable and Uncontrollable Aversive Events. *American Scientist, 60*(4), 457-465.

Gohier, B., Ferracci, L., Surguladze, S. A., Lawrence, E., El Hage, W., Kefi, M. Z., … Le Gall, D. (2009). Cognitive inhibition and working memory in unipolar depression. *Journal of Affective Disorders, 116*(1-2), 100–105. doi:10.1016/j.jad.2008.10.028

Goldberg, D. P., & Blackwell, B. (1970). Psychiatric Illness in General Practice: A Detailed Study Using a New Method of Case Identification. *BMJ, 2*(5707), 439–443. doi:10.1136/bmj.2.5707.439

Golden, C. J. (1978). *Stroop Color and Word Test: A Manual for Clinical and Experimental Uses*. Chicago, IL: Stoelting Co.

Golden, C. J., Espe-Pfeifer, P., & Wachsler-Felder, J. (2000) Neuropsychological interpretations of objective psychological tests. NY: Kluwer Academic/Plenum.

Grillon, C., Robinson, O. J., O'Connell, K., Davis, A., Alvarez, G., Pine, D. S., & Ernst, M. (2016). Clinical anxiety promotes excessive response inhibition. *Psychological Medicine, 47*(03), 484–494. doi:10.1017/s0033291716002555

Gådin, K. G., & Hammarström, A. (2000). School-Related Health—A Cross-Sectional Study among Young Boys and Girls. *International Journal of Health Services, 30*(4), 797–820. doi:10.2190/k3en-eay9-gdtd-002q

Hallion, L. S., Tolin, D. F., Assaf, M., Goethe, J., & Diefenbach, G. J. (2017). Cognitive Control in Generalized Anxiety Disorder: Relation of Inhibition Impairments to Worry and Anxiety Severity. *Cognitive Therapy and Research, 41*(4), 610–618. doi:10.1007/s10608-017-9832-2

Heo, M., Murphy, C. F., & Meyers, B. S. (2007). Relationship Between the Hamilton Depression Rating Scale and the Montgomery-Åsberg Depression Rating Scale in Depressed Elderly: A Meta-analysis. *The American Journal of Geriatric Psychiatry, 15*(10), 899–905. doi:10.1097/jgp.0b013e318098614e

Homack, S. (2004). A meta-analysis of the sensitivity and specificity of the Stroop Color and Word Test with children. *Archives of Clinical Neuropsychology, 19*(6), 725–743. doi:10.1016/j.acn.2003.09.003

Jensen, A. R. (1965). Scoring the Stroop Test. Acta Psychologia (24), 398 - 408. Scarpina, F., & Tagini, S. (2017). The Stroop Color and Word Test. *Frontiers in Psychology, 8*. doi:10.3389/fpsyg.2017.00557

Johnson, S. L. (2007). Cognitive inhibition across psychopathologies. *Applied and Preventive Psychology, 12*(3), 97–98. doi:10.1016/j.appsy.2007.11.001

Joormann, J., & Gotlib, I. H. (2010). Emotion regulation in depression: Relation to cognitive inhibition. *Cognition & Emotion, 24*(2), 281–298. doi:10.1080/02699930903407948

Joormann, J., Yoon, K. L., & Zetsche, U. (2007). Cognitive inhibition in depression. *Applied and Preventive Psychology, 12*(3), 128–139. doi:10.1016/j.appsy.2007.09.002

Kasch, K. L., Rottenberg, J., Arnow, B. A., & Gotlib, I. H. (2002). Behavioral activation and inhibition systems and the severity and course of depression. *Journal of Abnormal Psychology, 111*(4), 589–597. doi:10.1037/0021-843x.111.4.589

Kashyap, G. C., & Singh, S. K. (2017). Reliability and validity of general health questionnaire (GHQ-12) for male tannery workers: a study carried out in Kanpur, India. *BMC Psychiatry, 17*(1). doi:10.1186/s12888-017-1253-y

Kemeny, M. E. (2003). The Psychobiology of Stress. *Current Directions in Psychological Science, 12*(4), 124–129. doi:10.1111/1467-8721.01246

Kemeny, M. E., & Gruenewald, T. L. (2000). Affect, cognition, the immune system and health. *Progress in Brain Research*, 291–308. doi:10.1016/s0079-6123(08)62146-9

Kemeny, M. E., & Schedlowski, M. (2007). Understanding the interaction between psychosocial stress and immune-related diseases: A stepwise progression. *Brain, Behavior, and Immunity, 21*(8), 1009–1018. doi:10.1016/j.bbi.2007.07.010

Kivimäki, M., & Siegrist, J. (2016). Work Stress and Cardiovascular Disease: Reviewing Research Evidence with a Focus on Effort-Reward Imbalance at Work. In *Work Stress and Health in a Globalized Economy* (pp. 89–101). doi:10.1007/978-3-319-32937-6_5

Kivimäki, M., Virtanen, M., Elovainio, M., Kouvonen, A., Väänänen, A., & Vahtera, J. (2006). Work stress in the etiology of coronary heart disease—a meta-analysis. *Scandinavian Journal of Work, Environment & Health, 32*(6):431-442

Koolhaas, J. M., Bartolomucci, A., Buwalda, B., de Boer, S. F., Flügge, G., Korte, … Fuchs, E. (2011). Stress revisited: A critical evaluation of the stress concept. *Neuroscience & Biobehavioral Reviews, 35*(5), 1291–1301. doi:10.1016/j.neubiorev.2011.02.003

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9. *Journal of General Internal Medicine, 16*(9), 606–613. doi:10.1046/j.1525-1497.2001.016009606.x

Kroenke, K., Spitzer, R. L., Williams, J. B. W., Monahan, P. O., & Löwe, B. (2007). Anxiety Disorders in Primary Care: Prevalence, Impairment, Comorbidity, and Detection. *Annals of Internal Medicine, 146*(5), 317. doi:10.7326/0003-4819-146-5-200703060-00004

Li, C.-Y., & Sung, F.-C. (1999). A review of the healthy worker effect in occupational epidemiology. *Occupational Medicine, 49*(4), 225–229. doi:10.1093/occmed/49.4.225

Li, J. (2016). Working Conditions and Health in China: Evidence on the ERI Model. In *Work Stress and Health in a Globalized Economy* (pp. 219–234). doi:10.1007/978-3-319-32937-6_10

Li, J., Shang, L., Wang, T., & Siegrist, J. (2010). Measuring Effort–Reward Imbalance in School Settings: A Novel Approach and Its Association With Self-Rated Health. *Journal of Epidemiology, 20*(2), 111–118. doi:10.2188/jea.je20090057

Li, J., Yang, W., & Cho, S. (2006). Gender differences in job strain, effort-reward imbalance, and health functioning among Chinese physicians. *Social Science & Medicine, 62*(5), 1066–1077. doi:10.1016/j.socscimed.2005.07.011

MacLeod, C. (2015). The Stroop Effect. Encyclopedia of Color Science and Technology. doi: 10.1007/978-3-642-27851-8_67-1

MacLeod, C. M. (2007). The Concept of Inhibition in Cognition. In *Inhibition in Cognition* (pp. 3-23). Retrieved May 29, 2019, from https://pdfs.semanticscholar.org/a971/00c316e614f94dac998ac7fc750d4c4dd06c.pdf

Marková, J. (2018). *Stres a funkční poruchy žaludku* (Bachelor's thesis). Retrieved Jun. 20, 2019, from https://dspace.cuni.cz/bitstream/handle/20.500.11956/100871/130241167.pdf?sequence=1&isAllowed=y

Marmot, M. G., Smith, G. D., Stansfeld, S., Patel, C., North, F., Head, J., White, I., Brunner, E., Feeney, A. (1991). Health inequalities among British civil servants: the Whitehall II study. *The Lancet, 337*(8754), 1387–1393. doi:10.1016/0140-6736(91)93068-k

Marmot, M., Sapolsky, R. (2014). Of Baboons and Men: Social Circumstances, Biology, and the Social Gradient in Health. In *Sociality, Hierarchy, Health: Comparative Biodemography: A Collection of Papers*. Washington, D.C.: National Academies Press.

Martin, A., Rief, W., Klaiberg, A., & Braehler, E. (2006). Validity of the Brief Patient Health Questionnaire Mood Scale (PHQ-9) in the general population. *General Hospital Psychiatry, 28*(1), 71–77. doi:10.1016/j.genhosppsych.2005.07.003

Martin-Subero, M., Kroenke, K., Diez-Quevedo, C., Rangil, T., de Antonio, M., Morillas, R. M., … Navarro, R. (2017). Depression as Measured by PHQ-9 Versus Clinical Diagnosis as an Independent Predictor of Long-Term Mortality in a Prospective Cohort of Medical Inpatients. *Psychosomatic Medicine, 79*(3), 273–282. doi:10.1097/psy.0000000000000390

Minnich, A., Erford, B. T., Bardhoshi, G., Atalay, Z., Chang, C. Y., & Muller, L. A. (2019). Systematic Evaluation of Psychometric Characteristics of the Michigan Alcoholism Screening Test 13-Item Short (SMAST) and 10-Item Brief (BMAST) Versions. *Journal of Counseling & Development, 97*(1), 15–24. doi:10.1002/jcad.12231

Nigg, J. T. (2016). Annual Research Review: On the relations among self-regulation, self-control, executive functioning, effortful control, cognitive control, impulsivity, risk-taking, and inhibition for developmental psychopathology. *Journal of Child Psychology and Psychiatry, 58*(4), 361–383. doi:10.1111/jcpp.12675

Ortiz, V., G., & Juárez-García, A. (2016). Working Conditions and Effort-Reward Imbalance in Latin America. In *Work Stress and Health in a Globalized Economy* (pp. 235–271). doi:10.1007/978-3-319-32937-6_11

Ostry, A. S., Kelly, S., Demers, P. A., Mustard, C., & Hertzman, C. (2003). A comparison between the effort-reward imbalance and demand control models. *BMC Public Health, 3*(1). doi:10.1186/1471-2458-3-10

O'Malley, P. M. (2004). Maturing out of problematic alcohol use. *Alcohol Research & Health, 28*(4), 202-204.

Palmer, D. (2015). Cultural Change, Stress and Civil Servants' Occupational Health, c.1967–85. In *Stress in Post-War Britain, 1945–85* (pp. 95-109). New York, NY: Routledge

Patterson, S., Moran, P., Epel, E., Sinclair, E., Kemeny, M. E., Deeks, S. G., … Hecht, F. M. (2013). Cortisol Patterns Are Associated with T Cell Activation in HIV. *PLoS ONE, 8*(7), e63429. doi:10.1371/journal.pone.0063429

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., … Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods*. doi:10.3758/s13428-018-01193-y

Penner, I.-K., Kobel, M., Stöcklin, M., Weber, P., Opwis, K., & Calabrese, P. (2012). The Stroop Task: Comparison Between the Original Paradigm and Computerized Versions in Children and Adults. *The Clinical Neuropsychologist, 26*(7), 1142–1153. doi:10.1080/13854046.2012.713513

Pikhart, H., Bobák, M., Siegrist, J., Pajak, A., Rywik, S., Kyshegyi, J., … Marmot, M. (2001). Psychosocial Work Characteristics and Self Rated Health in Four Post-Communist Countries. *Journal of Epidemiology and Community Health, 55*(9), 624-630.

Pikhart, H., Bobak, M., Pajak, A., Malyutina, S., Kubinova, R., Topor, R., … Marmot, M. (2004). Psychosocial factors at work and depression in three countries of Central and Eastern Europe. *Social Science & Medicine, 58*(8), 1475–1482. doi:10.1016/s0277-9536(03)00350-2

Plummer, F., Manea, L., Trepel, D., & McMillan, D. (2016). Screening for anxiety disorders with the GAD-7 and GAD-2: a systematic review and diagnostic metaanalysis. *General Hospital Psychiatry, 39*, 24–31. doi:10.1016/j.genhosppsych.2015.11.005

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Richard-Devantoy, S., Jollant, F., Kefi, Z., Turecki, G., J.P.Olié, Annweiler, C., … Le Gall, D. (2012). Deficit of cognitive inhibition in depressed elderly: A neurocognitive marker of suicidal risk. *Journal of Affective Disorders, 140*(2), 193–199. doi:10.1016/j.jad.2012.03.006

Robert, M. A. & Hirschfeld, M. D. (2001). The Comorbidity of Major Depression and Anxiety Disorders:Recognition and Management in Primary Care. *Primary Care Companion to the Journal of Clinical Psychiatry, 3*(6), 244-254. doi:10.4088/pcc.v03n0609

Rugulies, R., Aust, B., & Madsen, I. E. H. (2016). Effort-Reward Imbalance and Affective Disorders. In *Work Stress and Health in a Globalized Economy* (pp. 145–170). doi:10.1007/978-3-319-32937-6_6

Ruiz, M. A., Zamorano, E., García-Campayo, J., Pardo, A., Freire, O., & Rejas, J. (2011). Validity of the GAD-7 scale as an outcome measure of disability in patients with generalized anxiety disorders in primary care. *Journal of Affective Disorders, 128*(3), 277–286. doi:10.1016/j.jad.2010.07.010

Sapolsky, R. M. (2012). Importance of a sense of control and the physiological benefits of leadership. *Proceedings of the National Academy of Sciences, 109*(44), 17730–17731. doi:10.1073/pnas.1215502109

Sapolsky, R. M. (2015). Stress and the brain: individual variability and the inverted-U. *Nature Neuroscience, 18*(10), 1344–1346. doi:10.1038/nn.4109

Scarpina, F., & Tagini, S. (2017). The Stroop Color and Word Test. *Frontiers in Psychology, 8*. doi:10.3389/fpsyg.2017.00557

Selzer, M. L., Vinokur, A., & van Rooijen, L. (1975). A self-administered Short Michigan Alcoholism Screening Test (SMAST). *Journal of Studies on Alcohol, 36*(1), 117–126. doi:10.15288/jsa.1975.36.117

Shang, L., Li, J., Li, Y., Wang, T., & Siegrist, J. (2013). Stressful psychosocial school environment and suicidal ideation in Chinese adolescents. *Social Psychiatry and Psychiatric Epidemiology, 49*(2), 205–210. doi:10.1007/s00127-013-0728-5

Shields, G. S., Sazma, M. A., & Yonelinas, A. P. (2016). The effects of acute stress on core executive functions: A meta-analysis and comparison with cortisol. *Neuroscience & Biobehavioral Reviews, 68*, 651–668. doi:10.1016/j.neubiorev.2016.06.038

Siegrist, J. (1996). Adverse health effects of high-effort/low-reward conditions. *Journal of Occupational Health Psychology, 1*(1), 27–41. doi:10.1037/1076-8998.1.1.27

Siegrist, J. (2008). Effort–reward imbalance and health in a globalized economy. *SJWEH Supplements*, (6), 163-168.

Siegrist, J. (2009). Unfair exchange and health: Social bases of stress-related diseases. *Social Theory & Health, 7*(4), 305–317. doi:10.1057/sth.2009.17

Siegrist, J. (2017). The Effort-Reward Imbalance Model. In *The Handbook of Stress and Health* (pp. 24–35). doi:10.1002/9781118993811.ch2

Siegrist, J., Starke, D., Chandola, T., Godin, I., Marmot, M., Niedhammer, I., & Peter, R. (2004). The measurement of effort–reward imbalance at work: European comparisons. *Social Science & Medicine, 58*(8), 1483–1499. doi:10.1016/s0277-9536(03)00351-4

Siegrist, M. (1997). Test-Retest Reliability of Different Versions of the Stroop Test. *The Journal of Psychology, 131*(3), 299–306. doi:10.1080/00223989709603516

Snyder, H. R. (2013). Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: A meta-analysis and review. *Psychological Bulletin, 139*(1), 81–132. doi:10.1037/a0028727

Sperlich, S., Arnhold-Kerri, S., Siegrist, J., & Geyer, S. (2012). The mismatch between high effort and low reward in household and family work predicts impaired health among mothers. *The European Journal of Public Health, 23*(5), 893–898. doi:10.1093/eurpub/cks134

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A Brief Measure for Assessing Generalized Anxiety Disorder. *Archives of Internal Medicine, 166*(10), 1092. doi:10.1001/archinte.166.10.1092

Starke, D., & Niedhammer, I. (2002). Gender and effort-reward imbalance at work. In S. P. Wamala & J. Lynch (Eds.), *Gender and social inequities in health*. Lund, Studentlitteratur

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643–662. doi:10.1037/h0054651

Svanborg, P., & Åsberg, M. (2001). A comparison between the Beck Depression Inventory (BDI) and the self-rating version of the Montgomery Åsberg Depression Rating Scale (MADRS). *Journal of Affective Disorders, 64*(2-3), 203–216. doi:10.1016/s0165-0327(00)00242-1

The British Psychological Society (2014). *Code of Human Research Ethics*. Retrieved Jun. 20, 2019, from https://www.bps.org.uk/sites/bps.org.uk/files/Policy/Policy - Files/BPS Code of Human Research Ethics.pdf

Tsutsumi, A. (2016). Work Stress and Health: The Case of Japan. In *Work Stress and Health in a Globalized Economy* (pp. 173-188). doi:10.1007/978-3-319-32937-6_8

Tsutsumi, A., Kayaba, K., Kario, K., & Ishikawa, S. (2009). Prospective Study on Occupational Stress and Risk of Stroke. *Archives of Internal Medicine, 169*(1), 56. doi:10.1001/archinternmed.2008.503

Van Vegchel, N. (2005). *Two models at work*. (Doctoral dissertation, Utrecht University, Utrecht, Netherlands). Retrieved Jun. 5, 2019, from https://dspace.library.uu.nl/handle/1874/7300

Van Vegchel, N., de Jonge, J., Bosma, H., & Schaufeli, W. (2005). Reviewing the effort–reward imbalance model: drawing up the balance of 45 empirical studies. *Social Science & Medicine, 60*(5), 1117–1131. doi:10.1016/j.socscimed.2004.06.043

Zdravotní Dotazník (PHQ-9). (n.d.). Retrieved Dec. 15, 2018, from https://www.phqscreeners.com/sites/g/files/g10049256/f/201412/PHQ9_Czech%20for% 20the%20Czech%20Republic.pdf