GROUP DIFFERENCES IN THE PROCESSING OF EMOTION-LADEN LINGUISTIC STIMULI

GROUP DIFFERENCES IN THE PROCESSING OF EMOTION-LADEN LINGUISTIC STIMULI

By: CONSTANCE L. IMBAULT, BSc. HON.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

McMaster University DOCTOR OF PHILOSOPHY (2019) Hamilton, Ontario (Cognitive Science of Language)

TITLE: GROUP DIFFERENCES IN THE PROCESSING OF EMOTION-LADEN LINGUISTIC STIMULI

AUTHOR: Constance L Imbault, BSc. Hon. (McMaster University)

SUPERVISOR: Drs. Victor Kuperman and David Shore

NUMBER OF PAGES: xi, 114

Abstract

This thesis investigated how different groups of people process the emotionality of linguistic stimuli. Recently, individual and group differences have become topics of interest in psychology, but very few studies have investigated how behavioural responses to emotional stimuli vary as a function of participant demographics.

The first study focused on the contrast between how native and non-native speakers of English process the emotional content of English words. The study of emotional processing in non-native speakers is a highly contested issue in the field of bilingualism and experimental psychology. Chapter 2 reports the largest collection of emotion ratings of English words from non-native speakers and provides a theoretical perspective on how bilingual speakers process emotion.

In order to facilitate the investigation of individual differences in emotional processing, the field needs a reliable method that allows for the measurement of subtle differences between individuals. In Chapter 3, I demonstrate that a novel 'slider' method of measuring valence, proposed by Warriner et al. (2018), is a reliable tool for measuring affective responses to words along a fine-grained sliding scale.

In Chapter 4, I demonstrate the use of the sliding scale with the aim of capturing affective differences between those with and without depressive symptoms. This chapter reports that compared to those without depressive symptoms, those with depressive symptoms i) exhibit attenuated responses to emotionally laden stimuli, and ii) are unable to take on the perspective of someone without depression.

Overall, this thesis reports on the emotional responses in two hotly debated groups, as well as providing a new method of measuring emotional responses to linguistic stimuli. These findings underscore the importance of studying emotional processing beyond normative populations.

ACKNOWLEDGEMENTS

I'd first and foremost like to thank my main supervisor, Victor Kuperman. Victor is an amazing supervisor; he's incredibly attentive, and is always there to bounce around ideas or to analyze new and exciting data. If it wasn't for him, I would not have been as productive as I have been in my PhD. He has such a passion for research that he shares with each of his students, and I hope this passion stays with me when I leave his lab. He gave me freedom to explore many topics that neither him nor I were an expert on, but which we shared a mutual curiosity. Victor always pushed me just outside of my comfort zone to allow me to grow both personally and as a researcher. He was there to cheer me on when I succeeded, but also there to teach me new lessons when I failed. I have learned an incredible amount from Victor and am so grateful that he allowed me to be a part of his lab.

How could I forget my other supervisor, David Shore. Although there were times where David forgot that I was one of his students, he always bounced back to being a great supervisor. He provided both laughs and support when I needed it most. He invited me to attend his cognition lab meetings, which he may not know how grateful I am for these meetings. Although these lab meetings may not have been directly applicable to my research, they allowed me to grow as a critical thinker, and be able to understand and dissect research where I am not an expert.

I am also very thankful for Louis Schmidt, the final member of my committee. Louis is always there to provide positive reassurance, and every interaction I have had with Louis has left me with a smile. He always had my best interests at heart and has given me wonderful, optimistic guidance.

I'd like to thank the Reading Lab for their support over the past 5 years. Many students have helped me along the way, running my studies and giving me constructive feedback. As well, I have formed great friendships in the lab that I hope will last a lifetime, wherever our hectic lives may take us.

Last, I'd like to thank Blake Anderson. He has always been there for me, talking me down from cliffs and patiently reassuring me when I didn't think I was enough. He has provided me with so much support, both emotionally and academically. "Blake Anderson has been very patient and helpful. I would not have persevered without his guidance throughout research, statistical analysis and the entire thesis process" (Imbault, 2014).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

DECLARATION OF ACADEMIC ACHIEVEMENT

The present dissertation constitutes a ``sandwich" thesis as defined by the School of Graduate Studies, McMaster University. I am the primary author of the three articles included in this dissertation, I conducted the literature reviews and wrote the manuscripts, designed the studies, and collected and analyzed all data. These studies comprised my doctoral research and are therefore included in the thesis. The roles of the co-authors for each paper are outlined below.

Chapter 2 is an article under review in *Bilingualism: Language and Cognition.*
Imbault, C., Warriner, A.B., Titone, D., & Kuperman, V. (Under Review). How are words felt in a second language: Norms for 2,628 English words for valence and arousal by non-native speakers

- Imbault, C.
  - Literature review; study design; data collection; analysis; manuscript writing, preparation, and revision
- Warriner, A.B.
  - Study design, data collection
- Titone, D.
  - Literature review, manuscript revisions
- Kuperman, V.
  - Study design; analysis, manuscript writing and revisions

Chapter 3 is a reprint of an article published in *Behavior Research Methods.*
Imbault, C., Shore, D., & Kuperman, V. (2018). Reliability of the sliding scale for collecting affective responses to words. *Behavior Research Methods*, 1–9. https://doi.org/10.3758/s13428-018-1016-9

- Imbault, C.
  - Literature review; study design; data collection; analysis; manuscript writing, preparation, and revision
- Shore, D.
  - Manuscript revisions
- Kuperman, V.
  - Study design; manuscript revisions

Chapter 4 is a reprint of an article published in *Scientific Reports.*

Imbault, C., & Kuperman, V. (2018). Emotional reactivity and perspective-taking in individuals with and without severe depressive symptoms. *Scientific Reports, 8*(1), 1–8. https://doi.org/10.1038/s41598-018-25708-x

- Imbault, C.
    - Literature review; study design; data collection; analysis; manuscript writing, preparation, and revision
- Kuperman, V.
    - Study design; analysis; manuscript revisions

Additional Achievements

In addition to what is presented in within the chapters of this thesis, the author was a primary contributor on 3 other studies.

Stadthagen-Gonzalez, H., Imbault, C., Pérez Sánchez, M. A., & Brysbaert, M. (2017). Norms of valence and arousal for 14,031 Spanish words. *Behavior Research Methods*, 1–45. https://doi.org/10.3758/s13428-015-0700-2

Stadthagen-González, H., Ferré, P., Pérez-Sánchez, M. A., Imbault, C., & Hinojosa, J. A. (2017). Norms for 10,491 Spanish words for five discrete emotions: Happiness, disgust, anger, fear, and sadness. *Behavior Research Methods*. https://doi.org/10.3758/s13428-017-0962-y

Warriner, A. B., Shore, D. I., Schmidt, L. A., Imbault, C. L., & Kuperman, V. (2017). Sliding into happiness: A new tool for measuring affective responses to words. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *71*(1), 71–88. https://doi.org/10.1037/cep0000112

Chapter 1

1. Introduction

1.1 Thesis Overview

Why do some people feel less emotional than others? And how do certain groups differ from others? The goal of this thesis is to investigate how different groups of people vary in how they process emotion. Despite emotion being a visceral, evolutionary response to stimuli, different groups of people can process the emotion of a stimulus completely differently. Individual differences methodologies and analyses are required to investigate these questions: record vast demographic details about our participants to understand which parts of their background affect how they process emotion. Individual differences have become of interest in psychology, and my goal was to apply an individual and group differences approach to emotional processing, specifically using linguistic stimuli. Furthermore, if there are differences in how different people process emotion, I explored whether these different groups are aware of the differences that exist between them.

Although intuitively straight-forward, a concrete and scientific definition of emotion has been hard to pin down. Kleinginna & Kleinginna (1981) collected 92 different definitions of emotion, which they split into 11 distinct categories. Most definitions contained a description of arousal and pleasure/displeasure (valence) and emphasized that emotion has many components, including affect, cognition, physiology and expressive behaviour. Furthermore, Kleinginna & Kleinginna (1981) created a distinction between affective definitions of emotion (categorizing emotions based on valence and arousal) and cognitive definitions of emotion (thinking about, appraising and labelling emotion), both of which are a primary focus of this thesis. They proposed a working definition of emotion:

> *Emotion is a complex set of interactions among subjective and objective factors, mediated by neural/hormonal systems, which can (a) give rise to affective experiences such as feelings of arousal, pleasure/displeasure; (b) generate cognitive processes such as emotionally relevant perceptual effects, appraisals, labeling processes; (c) activate widespread physiological*

adjustments to the arousing conditions; and (d) lead to behavior that is often, but not always, expressive, goal- directed, and adaptive.

There are two main theories that describe emotion. The Cannon-Bard theory of emotion describes emotion as a somatic response – the brain experiences a stimulus and feels an emotion, which results in a physiological reaction ("I am afraid, so I will tremble"). On the other hand, the James-Lange theory posits that an external stimulus results in a physiological response, and you feel an emotion because of the physiological response ("I am trembling, so I must be afraid"). Both theories agree that emotion is a mental state (whether it is caused or it causes a physiological reaction). The mental state of emotion gives rise to the differences in how emotion can be processed by different people.

Following our definitions of emotion, many researchers do not agree on how to break down affective categories. Some researchers prefer separating emotions into 5 discrete categories (anger, sadness, happiness, disgust and fear) (Ekman, 1992; Johnson-Laird & Oatley, 1989; Stadthagen-González, Ferré, Pérez-Sánchez, Imbault, & Hinojosa, 2017), while others prefer a bipolar, independent structure to define emotion (valence and arousal) (Stadthagen-Gonzalez, Imbault, Pérez Sánchez, & Brysbaert, 2017; Warriner, Kuperman, & Brysbaert, 2013). Valence is the bipolar positive to negative dimension, and arousal is the bipolar deactivation to activation dimension (Bradley & Lang, 1999; Feldman Barrett & Russell, 1998). Quantifying the number and labels of emotional dimensions is a hotly contested issue. Feldman Barrett & Russell (1998) found that valence and arousal were completely independent in what they measure. On the other hand, Warriner et al. (2013) found that valence and arousal are not entirely orthogonal dimensions, and rather, they have a U-shaped relationship. There is no clear-cut choice on how to measure emotion, but for the simplicity of using two dimensions, the entirety of this thesis will focus on valence and arousal.

The link between language and emotion may not be immediately obvious; language is a symbolic system without any inherent meaning other than what we have assigned to it. For example, a snake has an ecologically valid way of inducing a negative emotion (it is potentially a venomous creature), whereas the word "snake" or it phonological form /snejk/ is purely arbitrary. These three media of producing similar concepts have vastly different ecological validity, which is the extent to which reaction can be generalized into the real world (Wegener &

Blankenship, 2007). Although pictures and words convey similar concepts, a picture (or even better, a live snake) is a more realistic (and ecologically valid) way of communicating the idea than via language. We would expect more ecologically valid stimuli to induce greater emotional responses compared to less ecologically valid stimuli.

Although language and emotion are not ecologically linked, their processing systems are intertwined. Affective processing of linguistic stimuli is automatic. A classic behavioural example of this automaticity is with the Emotional Stroop Task (Dresler, Mériau, Heekeren, & Van Der Meer, 2009; Pratto & John, 1991; Williams, Mathews, & MacLeod, 1996), in which participants are given a list of words printed in different colours, and are told to report the colour of the word, whilst ignoring its meaning. Although told to solely focus on the colour, the emotion of the word affected their performance in the task; participants were slower to report the colour of emotional words compared to neutral words. This suggests that the affective processing of words is an aspect of cognitive behaviour that cannot be turned off.

Further examples of the automaticity of emotional processing of words comes from work on the physiological responses to linguistic stimuli. Emotion activates the peripheral nervous system (PNS), which is involved in eliciting the motivational “fight or flight” response. These responses guide the body into action following an arousing event (Rickard, 2004), preparing it to either fight or flee from danger. This reaction is evolutionary, automatic, and arises to both impending “real” threats, and linguistic stimuli representing a threat. Emotional words activate the PNS, causing the dilation of the pupils, increased heart and breathing rate and greater electrodermal activation, similar to how emotional pictorial stimuli would activate the PNS (Bayer, Sommer, & Schacht, 2011; De Pascalis, Fiore, & Sparita, 1996; De Pascalis, Strippoli, Riccardi, & Vergari, 2004; Võ et al., 2008).

Emotional words are processed differently than neutral words, giving further evidence for the relationship between linguistics and emotion. This topic has been of focus in the field for decades. Many studies have shown that emotional words are often named and recognized faster, and are more easily remembered, than neutral words (Altarriba & Bauer, 2004; Anooshian & Hertel, 1994; Bayer et al., 2011; Briesemeister, Kuchinke, & Jacobs, 2011; Kousta, Vinson, & Vigliocco, 2009; Kuperman, Estes, Brysbaert, & Warriner, 2014; Long & Titone, 2007; Mackay et al., 2004; Recio, Conrad, Hansen, & Jacobs, 2014; Scott, O’Donnell, Leuthold, & Sereno, 2009; N. A. Sheikh & Titone, 2013; Vinson, Ponari, & Vigliocco, 2014). These studies suggest that emotional words have a processing advantage over neutral words, and that emotional words may provide a more salient experience than neutral words.

There are several theories to explain how extra-linguistic stimuli, such as emotional connotation, are linked to our mental representation of a word. The embodiment account of word emotion proposes that the emotion of a word is learned through real-life, sensorimotor interactions with the concept (Barsalou, 1999; Clark & Paivio, 1991; Kousta, Vigliocco, Vinson, Andrews, & Del Campo, 2011). Barsalou (1999) theorizes that during an experience, the brain processes the experience in sensorimotor areas of the brain. During an emotional event, the emotional feelings are processed, and we later identify the specific parts of the experience that led to that emotion. This leads us to update our mental representation of the concept we experienced with specific emotional and sensorimotor information. This theory is in line with Dual Coding Theory (DCT; Clark & Paivio, 1991), which proposes that word representation is encoded not only by denotation, but also by our sensorimotor and emotional interactions with that word. Although they do not fully agree with DCT in terms of how it handles all concepts, Kousta et al. (2011) found that experiential information is bound to the meaning of linguistic concepts. It would seem that to have the full affective representation of a concept (whether linguistic, or otherwise), one would have to have experienced it and processed the concept in a sensorimotor way.

Given that the emotion of words is processed automatically, we can study the emotional processing of words as a means to investigate an individual's own responses to different emotional concepts. The first goal of this thesis was therefore to investigate the behavioural responses to emotional words by different groups of people. Individual and group differences are topics of interest in many fields, such as memory (Daneman & Carpenter, 1980; Just & Carpenter, 1999), second language learning (Ehrman, Leaver, & Oxford, 2003; Skehan, 1991) and reading (Kuperman & Van Dyke, 2011) to name a few. The recent focus on these differences has led to many advancements in these fields that were once overlooked at the population level.

In regards to emotional processing and individual differences, most studies have investigated their impact on neural activation. Eugene et al. (2003) explored the underpinnings as to why there is inconsistency in the literature about the neural correlates of sadness. They found individual differences in the regions of the brain that are activated when processing sad emotion (from film), which can explain past inconsistencies in the literature. Canli et al. (2001) found that an individual's personality traits affected which parts of the brain were activated while processing emotion. They found that those high in extraversion had greater activation during positive picture stimuli, and those high in neuroticism had greater neural activation during negative picture stimuli. These studies emphasize why the study of individual differences is so imperative; their findings were previously lost at the aggregate level of analysis. This need for individual differences is further stressed in another study by Canli et al. (2002), which found no difference in amygdala activation for happy stimuli at the aggregate level, but when analyzed at the individual level, they found a strong correlation between amygdala activation and an individual's extraversion score.

These findings were extended in a later behavioural study; Knyazev et al. (2007) found a wide range of personality traits that affected the processing of emotional faces. They found that increased behavioural inhibition and trait anxiety resulted in an increased rating of the hostility of faces; this effect was reversed for individuals high in intellect, agreeableness and conscientiousness.

In a study by Kuppens et al. (2017), both personality and cultural background affected the way individuals described and rated the valence and arousal of their own experiences. They found that extraverts rated their own experience as more arousing, and found their experiences to be more positive than introverts. People high in neuroticism tended to rate experiences as being more negative overall, those high in agreeableness tended to rate positive events as more arousing and negative events as less arousing. Kuppens et al. also investigated the effect of culture on how individuals described their own experiences. People from western cultures (Canada) tended to describe their experiences at extreme ends of the valence spectrum as highly arousing, whereas people from eastern cultures (Hong Kong) tended to describe the entire range of the valence spectrum as less arousing. This study shows the importance of collecting and analyzing participant variables, and the large impact they can have on results.

Less research has been done on the effects of individual or group differences on the emotional processing of linguistic stimuli. Bremner et al. (2001) found that women experience greater blood flow to the brain in response to emotional words than men. Additionally, women rated words as being more emotional than men, and remembered more emotional words than men did in a recall task. These both suggest that women have greater emotional responses to words than men. Despite these interesting findings, there seems to be little other research on the effects of individual and group differences on the processing of emotional linguistic stimuli. Warriner et al. (2018) found that those who are shyer than others tend to rate words as being more negative, but these findings were small and inconsistent. The goal of this thesis is to expand this work on group and individual differences to further sources of variation, from language background (monolingual vs. bilingual) to emotional background (those with and without depression).

There have been countless studies investigating how individual differences can affect that way we acquire and process emotion, yet very little research (discussed above) has been done to investigate how emotion and individual differences can play a role in language processing. Kidd, Donnelly, and Christiansen (2018) argued that we need to look at individual differences when creating models and theories of language acquisition and processing, such that we take into account the nature and environment where an individual learned the language. They also proposed that we need to think of our linguistic system as interconnected with other cognitive systems, such that they interact and affect each other. An individual's working memory capacity, statistical learning abilities, or their emotional reactivity, can greatly impact how they acquire and process language. The fact that individual differences in emotional reactivity has been largely ignored when creating linguistic models has severely impacted our understanding about how we learn and use language. The environment in which an infant is exposed to language can have effects that list into childhood (Weisleder & Fernald, 2013), and so it would make sense that an individual's emotional experience with particular concepts (often learned at young age) can also affect how they process those words later in life.

Designing studies, and separating participants based on individual difference measures have proven to be extremely useful in moving our knowledge forward. An important lesson to note is that of ecological fallacy. Ecological fallacy occurs when we interpret group data as a whole and deduce that nature of individuals from group-level data. A prime example of this occurring would be in a population with a bimodal (or more) distribution. The mean does not represent either group within the population, and if we deduce characteristics of the individuals from that population, we would be incorrectly drawing inferences about most individuals.

1.2 Non-native speaker's emotional processing

One group we investigate in this thesis is non-native speakers of English. English has the largest number of non-native speakers in the world (753.3 million, 2nd is Hindi with 274.2 million; Eberhard, Simons, & Fennig, 2009) and 14.9% of Canada's population have a first language other than English or French (Statistics Canada, 2013). This makes non-native speakers a particularly interesting and important group to study. Chapter 2 compares how and why non-native speakers of English process the emotion of English differently than native speakers.

There's a wealth of research on the differences in emotional processing between native and non-native speakers in a variety of languages. Many studies find that non-native speakers of a language do not experience the processing advantage of emotional words in the same way as native speakers (Anooshian & Hertel, 1994; Caldwell-Harris, Tong, Lung, & Poo, 2011; J. M. Dewaele & Nakano, 2013; Harris, Ayçiçeği, & Gleason, 2003; Santiago-Rivera & Altarriba, 2002; Sutton, Altarriba, Gianico, & Basnight-Brown, 2007). A non-native speaker may not find negative words as negative, or positive words as positive as a native speaker (extreme ends of the emotional spectrum become neutralized). For example, behaviourally, these come out as weakened effects of emotion on memory tasks (Anooshian & Hertel, 1994). Emotional words were better remembered in one's native language, but were recalled at similar levels to neutral words in one's non-native language.

Other studies, however, find that only one end of the emotional spectrum is impacted. Sheikh & Titone (2015) found that only positive words experienced a processing advantage while reading. Even more, other studies find no difference between native and non-native speakers (Ayçiçeği-Dinn & Caldwell-Harris, 2009; Eilola, Havelka, & Sharma, 2007; Ferré, García, Fraga, Sánchez-Casas, & Molero, 2010) or that this effect is attenuated with proficiency (Ponari, Rodriguez-Cuadrado, & Vinson, 2014) – the more proficient a non-native speaker becomes, the more native-like their emotional ratings become.

There are two main theories that drive this research. The proficiency account of these differences proposes that non-native speakers have a weaker lexical representation of words in their non-native language, but these lexical representations strengthen with proficiency (Perfetti, 2007). This theory proposes that because the lexical denotation of the word is weakened, non-native speakers simply don't respond as strongly to the emotion of these words. According to this theory, the more proficient a non-native speaker is in English, the more intact and native-like their emotional ratings should be. Furthermore, the more someone encounters a word, the greater the emotional bond. Thus, the word's own frequency (how often that word appears) should result in more native-like responses.

Other theories, discussed previously, propose that these differences are due to a lack of emotional grounding in a non-native speaker's life. For example, a non-native speaker may have learned (and become fluent in) a language through a book or within a classroom, without appropriate sensorimotor experiences associated with the language. Although they may understand the denotation of each word, they lack the emotional connotation. Barsalou (1999) proposes the emotion of language is built upon sensorimotor experiences; physical grounding of emotion and language in everyday life. According to this theory, words from the everyday environment where a non-native speaker uses English should be equally as emotional as a native speaker, but words from environments where a non-native speaker doesn't use English should experience emotional attenuation. For example, if a non-native speaker uses English at work, English words associated with work should receive native-like ratings. On the other hand, English words rarely used by non-native speakers (such as words relating to their family that they use at home) may experience greater emotional attenuation.

The goal of Chapter 2 was to investigate (1) if non-native speakers experience emotional words in English differently than native speakers, and (2) why there are differences between native and non-native speakers. Although there have been many studies that have looked into these differences, many of these studies have small sample sizes, reduced stimulus lists, and inconsistent or indirect ways of measuring the emotional processing of words. Additionally, no studies have taken into account the specific semantic categories of words to explain these differences in emotion perception. Chapter 2 rectifies these problems by providing a large dataset of words with a large number of participants. We also directly ask participants to rate the valence and arousal of words along a ratings scale (1-negative (least arousing), 9-positive (most arousing)), rather than measuring the behavioural or physiological responses to emotional words. As well, we examined the differences between native and non-native responses by looking at the effects of proficiency, length of time spent using English in Canada, as well as semantic categories of words. These can help us establish why non-native speakers have a different mental representation of emotional words in their non-native language compared to native speakers of the same language.

1.3 Reliability

In order to properly explore individual differences, the methods that are used need to be reliable. For a particular method to be reliable, it must produce similar results within an individual across multiple testing sessions. If a method is not reliable (ie. produces different results within the same individual), any differences between individuals may be due to inconsistencies within the method itself.

The goal of Chapter 3 is to test the reliability of a method that measures the valence (positivity) of a word. Warriner et al. (2018) designed a method that measures valence along a continuum. Traditional methods to measure valence typically involve asking participants to rate the valence of a word along a ratings scale (ex. 1 being negative to 9 being positive). These responses are then aggregated at the word level. A rating scale produces ordinal data but is often used and interpreted as interval data. For example, the difference between a rating of 2 and 3 is often interpreted to be the same as the difference between 3 and 4, but this is not necessarily true along a rating scale. Additionally, the measurement from a typical rating scale is too course-grained to measure subtle individual differences (C. J. Russell & Bobko, 1992). To remedy these problems, Warriner et al. (2018) proposed a sliding scale, which allows participants to choose their response along a 648-point continuum. The sliding scale is presented on a computer screen with a humanoid manikin (a cartoon person who the participants are instructed represents themselves). Target words are displayed above or below the manikin, and participants are instructed to slide the manikin closer to or further away from the word depending on their own emotional response towards the word. The manikin's final distance from the word (measured in pixels, 1-648 pixels) represents the participant's valence response; a smaller distance (close) indicates greater positivity, a larger distance (far) indicates negativity. We compare the manikin's final distance with more traditional measures of valence (ratings scale) to ensure the sliding scale is measuring valence.

Furthermore, another benefit of this method, over traditional rating scales, is that it can tap into evolutionary approach-avoidance behaviours. Approach motivation is the direction of behaviour toward positive stimuli, whereas avoidance behaviour is the direction of behaviour away from negative stimuli (Elliot, 2006). This behaviour is automatic, and we evaluate all stimuli based on this positive/negative dimension (Chen & Bargh, 1997). The sliding scale was tested for validity and usability in a recent paper by Warriner et al. (2018). They found a consistent relationship between valence and sliding distance, and were able to capture individual differences between genders and across participant shyness scores.

We tested the reliability of the sliding scale with three separate measures: split-half reliability, one session reliability, and two session reliability. Split-half reliability measures the internal consistency of a measure (Macleod et al., 2010). To measure split-half reliability, the stimulus list is randomly split into two halves. We then find the effect of valence on distance (the final position of the manikin) for each participant across both halves of the stimuli. We measure the correlation between the two halves to find a single measure of reliability. This method is repeated 1000 times to get 1000 measures of reliability, and we then average these correlations coefficients for the split-half reliability and calculate 95% confidence intervals.

One session reliability measures the internal consistency of a measure across a single session. To do so, we present participants with two blocks of the same data. We then correlate the ratings from first timepoint with the ratings from the second timepoint to find the consistency across the same session. Two session reliability performs a similar function to one session reliability, but instead extends the reliability across a span of time (typically a day to a week). Participants perform the task at the first timepoint, and respond to the same words at a second timepoint. Similar to one session reliability, we calculate the consistency across the two sessions.

With evidence that the sliding scale is a reliable measure, we can use it in ways to measure subtle individual differences that may have been lost with a coarser measurement of valence. Furthermore, with the slider task being implemented through a computer program, the appearance of the slider and the manikin can be altered to answer different research questions from different groups.

1.4 Perspective Taking

Another interesting use of the slider paradigm is the ability to change the picture of the manikin. In the typical slider task, a humanoid manikin is used to represent the participant. The participant moves the manikin as close to or as far away from a word as they want, and their final position (measured in pixels) is representative of how they would rate the valence of that word. But, because the slider is an online task, the image of the humanoid manikin can change throughout the experiment. This means that we can use the slider task to elicit responses for someone other than the participant themselves; we can ask the participant to take on the perspective of another person.

Perspective-taking tasks typically ask the participant to imagine how another person feels in a given situation, and then observe how they respond to stimuli or questionnaires after the perspective taking task (Batson, Early, & Salvarani, 1997). In other studies, perspective-taking skill is measured through a questionnaire rather than a task (Schutte et al., 2001), and their perspective-taking skill (from the questionnaire) is correlated with their performance in another task. These tasks test the perspective taking ability of an individual, but they do not capture the participants' predictions of how the other person behaves.

The slider task allows for a more intuitive, direct way to measure both perspective taking abilities, but also how the participants predict others would respond. With perspective taking instructions, we can compare predicted behaviour of an individual to their actual behaviour.

1.5 Depression

The goal of Chapter 4 was two-fold; first to measure the emotional reactivity of those with and without depressive symptoms, and second to investigate whether they can predict how the opposite groups feels. 11.3% of adults in Canada will meet the criteria for depression in their lifetime (Pearson, Janz, & Ali, 2013), which makes them an important group to study. Additionally, people with depression are particularly relevant to the study of emotion as one of the diagnostic criteria for depression is defined as the lack of pleasure or happiness (American Psychiatric Association, 2013). Alternatively, those with depression sometimes describe feeling "numb". Although these two descriptions may seem quite similar, loss of pleasure is neutralization in the positive domain (positive items are no longer positive), while "numb" mood may exhibit as neutralization over both the positive and negative domain. These two versions of emotional reactivity have been hotly debated in the literature. Some theorize that depression is negative potentiation of mood (pervasive negative mood) (Golin, Hartman, Klatt, Munz, & Wolfgang, 1977; Lewinsohn, Lobitz, & Wilson, 1973; Sigmon & Nelson-Gray, 1992), while others theorize positive attenuation (lack of positive mood) (Allen, Trinder, & Brennan, 1999; Sloan, Strauss, Quirk, & Sajatovic, 1997; Sloan, Strauss, & Wisner, 2001). Another camp suggests that those with depression experience emotion context insensitivity, the attenuation of emotional reactivity across both positive and negative domains (Rottenberg, Gross, & Gotlib, 2005; Rottenberg & Hindash, 2015). Using the slider paradigm, we can record explicit emotional responses to words, and by doing so, can investigate the emotional reactivity of those with depression. Bylsema et al. (2007) performed a meta-analysis of the depression literature and found that across many studies, both positive and negative emotional reactivity were reduced – most studies found emotion context insensitivity in participants with depression.

The second question in Chapter 4 is whether those with and without depressive symptoms are aware of the other's emotional reactivity. The fact that the emotional reactivity of those with depression is hotly debated within the scientific community leads us to believe that lay participants may be unable to predict the emotional reactivity of someone with depression. Being understood, and feeling a sense of community and support is value for the prognosis of depression. Those who report stronger social networks experienced lower levels of depression (Cutrona & Troutman, 1986; Peirce, Frone, Russell, Cooper, & Mudar, 2000). If our lay participants are able to take on the perspective of someone with depression, it may lead to a greater sense of community and understanding for those with depression.

Furthermore, another interesting outcome of this manipulation is whether those with depressive symptoms can understand what it feels like to not have depression. The hopelessness theory of depression states that in depression, people may feel despondent (Abramson, Metalsky, & Alloy, 1989). If you're unable to understand what it is like to not feel depressed, autonomous motivation can be lost. Autonomous motivation, or the ability to want to get better, is a strong predictor of lower remission and depression severity (Zuroff et al., 2007). Chapter 4 hopes to display the emotional reactivity of those with depressive symptoms, as well as uncover whether those with and without depressive symptoms can understand and predict how the other group feels.

1.6 Conclusion

All in all, the goal of this thesis is to push the investigation of individual and group differences forward by providing evidence for individual differences in emotional processing, as well as providing a suitable method for researchers to forward this line of research. Many interesting findings have been lost when studying people at the aggregate level, and I believe that we can uncover many important and exciting differences between different groups of people. This thesis should serve as motivation to continue investigating the effects of individual and group differences at all levels.

References

Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review*, *96*(2), 358–372. https://doi.org/10.1037/0033-295X.96.2.358

Albaum, G., Best, R., & Hawkins, D. E. L. (1981). Continuous vs Discrete Semantic Differential Rating Scales. *Psychological Reports*, *49*, 83–86.

Allen, N. B., Trinder, J., & Brennan, C. (1999). Affective Startle Modulation in Clinical Depression: Preliminary Findings. *Biological Psychiatry*, *46*(4), 542–550. https://doi.org/https://doi.org/10.1016/S0006-3223(99)00025-6

Altarriba, J., & Bauer, L. M. (2004). The Distinctiveness of Emotion Concepts: A Comparison between Emotion, Abstract, and Concrete Words. *The American Journal of Psychology*, *117*(3), 389–410.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (DSM-5®). *American Psychiatric Pub*.

Anooshian, L. J., & Hertel, P. T. (1994). Emotionality in free recall: Language specificity in bilingual memory. *Cognition and Emotion*, *8*(6), 503–514. https://doi.org/10.1080/02699939408408956

Ayçiçeği-Dinn, A., & Caldwell-Harris, C. L. (2009). Emotion-memory effects in bilingual speakers: A levels-of-processing approach*. *Bilingualism: Language and Cognition*, *12*(03), 291. https://doi.org/10.1017/S1366728909990125

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(04), 577–660. https://doi.org/10.1017/S0140525X99002149

Batson, C. D., Early, S., & Salvarani, G. (1997). Perspective Taking: Imagining How Another Feels Versus Imaging How You Would Feel. *Personality and Social Psychology Bulletin*. https://doi.org/10.1177/0146167297237008

Bayer, M., Sommer, W., & Schacht, A. (2011). Emotional words impact the mind but not the body: Evidence from pupillary responses. *Psychophysiology*, *48*(11), 1554–1562. https://doi.org/10.1111/j.1469-8986.2011.01219.x

Betella, A., & Verschure, P. F. M. J. (2016). The affective slider: A digital self-assessment scale for the measurement of human emotions. *PLoS ONE*, *11*(2), 1–11. https://doi.org/10.1371/journal.pone.0148037

Bonin, P., Aubert, L., Malardier, N., & Niedenthal, P. M. (2016). Subjective Et De Valence Émotionnelle Pour 866 Mots, *103*, 655–694.

Bradley, M. M., & Lang, P. P. J. (1999). Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings. *Psychology*, *Technical*(C–1), 0. https://doi.org/10.1109/MIC.2008.114

Brase, J., & Mani, N. (2017). Effects of learning context on the acquisition and processing of emotional words in bilinguals. *Emotion*, *17*(4), 628–639. https://doi.org/10.1037/emo0000263

Briesemeister, B. B., Kuchinke, L., & Jacobs, A. M. (2011). Discrete emotion effects on lexical decision response times. *PLoS ONE*, *6*(8). https://doi.org/10.1371/journal.pone.0023743

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990. https://doi.org/10.3758/BRM.41.4.977

Bylsma, L. M., Morris, B. H., & Rottenberg, J. (2008). A meta-analysis of emotional reactivity in major depressive disorder. *Clinical Psychology Review*, *28*(4), 676–691. https://doi.org/10.1016/j.cpr.2007.10.001

Caldwell-Harris, C. L. (2015). Emotionality Differences Between a Native and Foreign Language. *Current Directions in Psychological Science*, *24*(3), 214–219. https://doi.org/10.1177/0963721414566268

Caldwell-Harris, C. L., Tong, J., Lung, W., & Poo, S. (2011). Physiological reactivity to emotional phrases in Mandarin--English bilinguals. *International Journal of Bilingualism*, *15*(3), 329–352. https://doi.org/10.1177/1367006910379262

Canli, T., Sivers, H., Whitfield, S. L., Gotlib, I. H., & Gabrieli, J. D. E. (2002). Amygdala response to happy faces as a function of extraversion. *Science*, *296*(5576), 2191. https://doi.org/10.1126/science.1068749

Cartwright, C., Gibson, K., Read, J., Cowan, O., & Dehar, T. (2016). Long-term antidepressant use: patient perspectives of benefits and adverse effects. *Patient Preference and Adherence*, *Volume 10*, 1401–1407. https://doi.org/10.2147/PPA.S110632

Carver, C. S., & White, T. L. (1994). Behavioral Inhibition, Behavioral Activation and affective responses to impending reward and punishment. *Journal of Personality and Social Psychology*, *67*, 319–333.

Chen, M., & Bargh, J. A. (1997). Nonconscious behavioral confirmation processes: The self-fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology*, *33*(5), 541–560. https://doi.org/10.1006/jesp.1997.1329

Citron, F. M. M., Weekes, B. S., & Ferstl, E. C. (2013). Effects of valence and arousal on written word recognition: Time course and ERP correlates. *Neuroscience Letters*, *533*, 90–95.

Clark, J., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*(3), 149–210. https://doi.org/10.1007/BF01320076

Cutrona, C. E., & Troutman, B. R. (1986). Social Support , Infant Temperament , and Parenting Self-Efficacy : A Mediational Model of Postpartum Depression Author (s): Carolyn E . Cutrona and Beth R . Troutman Published by : Wiley on behalf of the Society for Research in Child Development Stable, *57*(6), 1507–1518.

Daneman, M., & Carpenter, P. A. (1980). Individual Differences in Working Memory and Reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466.

De Grauwe, S., Willems, R. M., Rueschemeyer, S. A., Lemhöfer, K., & Schriefers, H. (2014). Embodied language in first- and second-language speakers: Neural correlates of processing motor verbs. *Neuropsychologia*, *56*(1), 334–349. https://doi.org/10.1016/j.neuropsychologia.2014.02.003

De Pascalis, V., Fiore, A. D., & Sparita, A. (1996). Personality, event-related potential (ERP) and heart rate (HR): An investigation of Gray's theory. *Personality and Individual Differences*, *20*(6), 733–746. https://doi.org/10.1016/0191-8869(96)00016-5

De Pascalis, V., Strippoli, E., Riccardi, P., & Vergari, F. (2004). Personality, event-related potential (ERP) and heart rate (HR) in emotional word processing. *Personality and Individual Differences*, *36*(4), 873–891. https://doi.org/10.1016/S0191-8869(03)00159-4

Decety, J., & Batson, C. D. (2007). Social neuroscience approaches to interpersonal sensitivity. *Social Neuroscience*, *2*(3–4), 151–157. https://doi.org/10.1080/17470910701506060

Degner, J., Doycheva, C., & Wentura, D. (2012). It matters how much you talk: On the automaticity of affective connotations of first and second language words. *Bilingualism*, *15*(1), 181–189. https://doi.org/10.1017/S1366728911000095

Dewaele, J.-M. (2004). The Emotional Force of Swearwords and Taboo Words in the Speech of Multilinguals. *Journal of Multilingual and Multicultural Development*, *25*(23), 204–222. https://doi.org/10.1080/01434630408666529

Dewaele, J. M., & Nakano, S. (2013). Multilinguals' perceptions of feeling different when switching languages. *Journal of Multilingual and Multicultural Development*, *34*(2), 107–120. https://doi.org/10.1080/01434632.2012.712133

Dresler, T., Mériau, K., Heekeren, H. R., & Van Der Meer, E. (2009). Emotional Stroop task: Effect of word arousal and subject anxiety on emotional interference. *Psychological Research*, *73*(3), 364–371. https://doi.org/10.1007/s00426-008-0154-6

Ehrman, M. E., Leaver, B. Lou, & Oxford, R. L. (2003). A brief overview of individual differences in second language learning. *System*, *31*(3), 313–330. https://doi.org/10.1016/S0346-251X(03)00045-9

Eilola, T. M., Havelka, J., & Sharma, D. (2007). Emotional activation in the first and second language. *Cognition and Emotion*, *21*(5), 1064–1076. https://doi.org/10.1080/02699930601054109

Eisman, A. B., Stoddard, S. a, Heinze, J., Caldwell, C. H., & Zimmerman, M. a. (2015). Depressive symptoms, social support, and violence exposure among urban youth: A longitudinal study of resilience. *Developmental Psychology*, *51*(9), 1307–1316. https://doi.org/10.1037/a0039501

Ekman, P. (1992). An Argument for Basic Emotions. *Cognition and Emotion*, *6*(3–4), 169–200. https://doi.org/10.1080/02699939208411068

Elliot, A. J. (2006). The Hierarchical Model of Approach-Avoidance Motivation. *Motivation and Emotion*, *30*(2), 111–116. https://doi.org/10.1007/s11031-006-9028-7

Epstein, J., Pan, H., Kocsis, J., Yang, Y., Butler, T., Chusid, J., … Silbersweig, D. (2006). Lack of Ventral Striatal Response to Positive Stimuli inDepressed Versus Normal Subjects. *The American Journal of Psychiatry*, 1784. https://doi.org/10.1176/ajp.2006.163.10.1784

Farrow, T. F. D., & Woodruff, P. W. R. (2008). Empathy in Mental Illness. *Psychiatric Services*, *59*, 581–581. https://doi.org/10.1176/appi.ps.59.5.581

Feldman Barrett, L., & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology*, *74*(4), 967–984. https://doi.org/10.1037/0022-3514.74.4.967

Ferré, P., García, T., Fraga, I., Sánchez-Casas, R., & Molero, M. (2010). Memory for emotional words in bilinguals: Do words have the same emotional intensity in the first and in the second language? *Cognition and Emotion*, *24*(5), 760–785. https://doi.org/10.1080/02699930902985779

Foroni, F. (2015). Do we embody second language? Evidence for ‘partial’ simulation during processing of a second language. *Brain and Cognition*, *99*, 8–16. https://doi.org/10.1016/j.bandc.2015.06.006

George, L., Blazer, D., Hughes, D., & Fowler, N. (1989). Social support and the outcome of major depression. *British Journal of Psychiatry*, *154*, 478–485.

Golin, S., Hartman, S. A., Klatt, E. N., Munz, K., & Wolfgang, G. L. (1977). Effects of Self-esteem Manipulation on Arousal and Reactions to Sad Models in Depressed and Nondepressed College Students. *Journal of Abnormal Psychology*, *86*(4), 435–439.

Haaga, D. A. F., & Beck, A. T. (1995). Perspectives on depressive realism: Implications for cognitive theory of depression. *Behaviour Research and Therapy*, *33*(1), 41–48. https://doi.org/10.1016/0005-7967(94)E0016-C

Harris, C. L., Ayçiçeği, A., & Gleason, J. B. (2003). Taboo words and reprimands elicit greater autonomic reactivity in a first language than in a second language. *Applied Psycholinguistics*, *24*(4), 561–579. https://doi.org/10.1017/S0142716403000286

Harris, C. L., Gleason, J. B., & Ayçiçeği, A. (2006). When is a First Language More Emotional? Psychophysiological Evidence from Bilingual Speakers. In *Bilingual minds: Emotional experience, expression and representation* (pp. 257–283). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.69.5830%0Ahttp://www.zeit.de/wirtschaft/2014-08/russland-import-sanktionen%0Ahttp://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.69.5830%0Ahttp://www.zeit.de/wirtschaft/2014-08/russland-import-sa

Hellman, A. B. (2011). Vocabulary size and depth of word knowledge in adult-onset second language acquisition. *International Journal of Applied Linguistics*, *21*(2), 162–182. https://doi.org/10.1111/j.1473-4192.2010.00265.x

Hinojosa, J. a., Martínez-García, N., Villalba-García, C., Fernández-Folgueiras, U., Sánchez-Carmona, a., Pozo, M. a., & Montoro, P. R. (2015). Affective norms of 875 Spanish words for five discrete emotional categories and two emotional dimensions. *Behavior Research Methods*, (August). https://doi.org/10.3758/s13428-015-0572-5

Imbault, C., Shore, D., & Kuperman, V. (2017). Reliability of the sliding scale for collecting affective responses to words. *Behavior Research Methods*.

Inoue, Y., Yamada, K., & Kanba, S. (2006). Deficit in theory of mind is a risk for relapse of major depression. *Journal of Affective Disorders*, *95*(1–3), 125–127. https://doi.org/10.1016/j.jad.2006.04.018

Inoue, Yumiko, Tonooka, Y., Yamada, K., & Kanba, S. (2004). Deficiency of theory of mind in patients with remitted mood disorder. *Journal of Affective Disorders*, *82*(3), 403–409. https://doi.org/10.1016/j.jad.2004.04.004

Itkes, O., Eviatar, Z., & Kron, A. (2019). Semantic and affective manifestations of ambi (valence). *Cognition and Emotion*, *9931*. https://doi.org/10.1080/02699931.2018.1564249

Itkes, O., Kimchi, R., Kron, A., & Carmel, M. (2017). Dissociating Affective and Semantic Valence, *146*(7), 1–36. https://doi.org/10.1037/xge0000291

Itzhak, I., Vingron, N., Baum, S. R., & Titone, D. (2017). Bilingualism in the real world: How proficiency, emotion, and personality in a second language impact communication in clinical and legal settings. *Translational Issues in Psychological Science*, *3*(1), 48–65. https://doi.org/10.1037/tps0000103

Johnson-Laird, P. N., & Oatley, K. (1989). The Language of Emotions: An Analysis of a Semantic Field. *Cognition and Emotion*, *3*(2), 81–123. https://doi.org/10.1080/02699938908408075

Jolliffe, D., & Farrington, D. P. (2006). Development and validation of the Basic Empathy Scale. *Journal of Adolescence*, *29*(4), 589–611. https://doi.org/10.1016/j.adolescence.2005.08.010

Just, M. A., & Carpenter, P. A. (1999). A Capacity Theory of Comprehension : Indicidual Dirrefences in Working Memory, (January), 1–16.

Kazanas, S. A., & Altarriba, J. (2016). Emotion Word Processing: Effects of Word Type and Valence in Spanish–English Bilinguals. *Journal of Psycholinguistic Research*, *45*(2), 395–406. https://doi.org/10.1007/s10936-015-9357-3

Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual Differences in Language Acquisition and Processing. *Trends in Cognitive Sciences*, *22*(2), 154–169. https://doi.org/10.1016/j.tics.2017.11.006

Kleinginna, P. R., & Kleinginna, A. M. (1981). A categorized list of motivation definitions, with a suggestion for a consensual definition. *Motivation and Emotion*, *5*(3), 263–291. https://doi.org/10.1007/BF00993889

Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The Representation of Abstract Words: Why Emotion Matters. *Journal of Experimental Psychology: General*, *140*(1), 14–34. https://doi.org/10.1037/a0021446

Kousta, S. T., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition*, *112*(3), 473–481. https://doi.org/10.1016/j.cognition.2009.06.007

Kuchinke, L., Võ, M. L. H., Hofmann, M., & Jacobs, A. M. (2007). Pupillary responses during lexical decisions vary with word frequency but not emotional valence. *International Journal of Psychophysiology*, *65*(2), 132–140. https://doi.org/10.1016/j.ijpsycho.2007.04.004

Kunzmann, U., & Grühn, D. (2005). Age differences in emotional reactivity: The sample case of sadness. *Psychology and Aging*, *20*(1), 47–59. https://doi.org/10.1037/0882-7974.20.1.47

Kuperman, V., Estes, Z., Brysbaert, M., & Warriner, A. B. (2014). Emotion and language: Valence and arousal affect word recognition. *Journal of Experimental …*, (4). Retrieved from http://psycnet.apa.org/psycarticles/2014-03909-001

Kuperman, V., & Van Dyke, J. a. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language*, *65*(1), 42–73. https://doi.org/10.1016/j.jml.2011.03.002

Kuppens, P., Tuerlinckx, F., Yik, M., Koval, P., Coosemans, J., Zeng, K. J., & Russell, J. A. (2017). The Relation Between Valence and Arousal in Subjective Experience Varies With Personality and Culture. *Journal of Personality*, *85*(4), 530–542. https://doi.org/10.1111/jopy.12258

Kuznetsova, A., Brockhoff, P., & Christensen, R. (2016). lmerTest: Tests in Linear Mixed Effects Models. *R Package Version*, *3.0.0*, https://cran.r-project.org/package=lmerTest. Retrieved from https://cran.r-project.org/package=lmerTest

Lang, P. J. (1995). The Emotion Probe: Studies of Motivation and Attention. *American Psychologist*. https://doi.org/10.1037/0003-066X.50.5.372

Larsen, R. J., & Ketelaar, T. (1991). Personality and susceptibility to positive and negative emotional states. *Journal of Personality and Social Psychology*, *61*(1), 132–140. https://doi.org/10.1037/0022-3514.61.1.132

Levenson, R. W., Carstensen, L. L., Friesen, W. V., & Ekman, P. (1991). Emotion, physiology, and expression in old age. *Psychology and Aging*, *6*(1), 28–35. https://doi.org/10.1037/0882-7974.6.1.28

Lewinsohn, P. M., Lobitz, W. C., & Wilson, S. (1973). “Sensitivity” of Depressed Individuals to Aversive Stimuli. *Journal of Abnormal Psychology*, *81*(3), 259–263.

Long, L. S., & Titone, D. A. (2007). Psychopathy and verbal emotion processing in non-incarcerated males. *Cognition and Emotion*, *21*(1), 119–145. https://doi.org/10.1080/02699930600551766

Mackay, D. G., Shafto, M., Taylor, J. K., Marian, D. E., Abrams, L., & Dyer, J. R. (2004). Relations between emotion, memory, and attention: Evidence from taboo Stroop, lexical decision, and immediate memory tasks. *Memory & Cognition*, *32*(3), 474–488. https://doi.org/10.3758/BF03195840

Macleod, J. W., Lawrence, M. a, McConnell, M. M., Eskes, G. a, Klein, R. M., & Shore, D. I. (2010). Appraising the ANT: Psychometric and theoretical considerations of the Attention Network Test. *Neuropsychology*, *24*(5), 637–651. https://doi.org/10.1037/a0019803

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research : JSLHR*, *50*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

Mcwhirter, B. T., Besett-Alesch, T. M., Horibata, J., & Gat, I. (2002). Loneliness in High Risk Adolescents: The Role of Coping, Self-Esteem, and Empathy. *Journal of Youth Studies*, *5*(1), 69–84. https://doi.org/10.1080/1367626012011177

Méndez-Bértolo, C., Pozo, M. A., & Hinojosa, J. A. (2011). Word frequency modulates the processing of emotional words: Convergent behavioral and electrophysiological data. *Neuroscience Letters*, *494*(3), 250–254. https://doi.org/10.1016/j.neulet.2011.03.026

Monnier, C., & Syssau, A. (2014). Affective Norms for French Words (FAN). *Behavior Research Methods*, *46*(4), 1128–1137. https://doi.org/10.3758/s13428-013-0431-1

Moors, A., De Houwer, J., Hermans, D., Wanmaker, S., van Schie, K., Van Harmelen, A.-L., … Brysbaert, M. (2013). Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words. *Behavior Research Methods*, *45*(1), 169–177. https://doi.org/10.3758/s13428-012-0243-8

Mor, N., & Winquist, J. (2002). Self-focused attention and negative affect: A meta-analysis. *Psychological Bulletin*, *128*(4), 638–662. https://doi.org/10.1037/0033-2909.128.4.638

O'Connor, L. E., Berry, J. W., Lewis, T., Mulherin, K., & Crisotomo, P. S. (2007). Empathy and depression: the moral system on overdrive. *Empathy in Mental Ilness*. https://doi.org/10.1017/CBO9780511543753.005

Orne, M. T. (1962). On the Social Psychology of Psychological Experiment: With Particular Reference to Demand Characteristics and Their Implications. *American Psychologist, 17*(11), 776–783. Retrieved from https://journals.scholarsportal.info/details/0003066x/v17i0011/776_otspottdcati.xml

Ostaszewski, K., & Zimmerman, M. A. (2006). The effects of cumulative risks and promotive factors on urban adolescent alcohol and other drug use: A longitudinal study of resiliency. *American Journal of Community Psychology*, *38*(3–4), 237–249. https://doi.org/10.1007/s10464-006-9076-x

Pearson, C., Janz, T., & Ali, J. (2013). *Mental and substance use disorders in Canada*.

Peirce, R., Frone, M., Russell, M., Cooper, M., & Mudar, P. (2000). A Longitudinal Model of Social Contact, Social Support, Depression, and Alcohol Use. *Health Psychol*, *19*(1), 28–38. https://doi.org/10.1037/0278-6133.19.1.28

Pennebaker, J., Booth, R., & Francis, M. (2007). Linguistic inquiry and word count: LIWC. Austin, TX: liwc net.

Perfetti, C. (2007). Reading Ability: Lexical Quality to Comprehension. *Scientific Studies of Reading*, *11*(4), 357–383. https://doi.org/10.1080/10888430701530730

Ponari, M., Rodriguez-Cuadrado, S., & Vinson, D. (2014). Processing advantage for emotional words in bilingual speakers. *Emotion*, *15*(January), 644–652. https://doi.org/10.1037/emo0000061

Ponari, M., Rodriguez-Cuadrado, S., Vinson, D., Fox, N., Costa, A., & Vigliocco, G. (2014). Processing advantage for emotional words in bilingual speakers. *Emotion*, *15*(January), 644–652. https://doi.org/10.1037/emo0000061

Pratto, F., & John, O. P. (1991). Automatic Vigilance: The Attention-Grabbing Power of Negative Social Information. *Journal of Personality and Social Psychology*, *61*(3), 380–391. https://doi.org/10.1037/0022-3514.61.3.380

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *The Behabioral and Brain Sciences*, *4*, 515–526. https://doi.org/10.1016/j.celrep.2011.1011.1001.7.

Puntoni, S., de Langhe, B., & van Osselaer, S. M. J. (2009). Bilingualism and the Emotional Intensity of Advertising Language. *Journal of Consumer Research*, *35*(6), 1012–1025. https://doi.org/10.1086/595022

Pyszczynski, T., Holt, K., & Greenberg, J. (1987). Depression, self-focused attention, and expectancies for positive and negative future life events for self and others. *Journal of Personality and Social Psychology*, *52*, 994–1001. https://doi.org/10.1037/0022-3514.52.5.994

R Core Team, I. (2017). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna (2014). R Foundation for Statistical Computing Authors R Core Team.

Recio, G., Conrad, M., Hansen, L. B., & Jacobs, A. M. (2014). On pleasure and thrill: The interplay between arousal and valence during visual word recognition. *Brain and Language*, *134*, 34–43. https://doi.org/10.1016/j.bandl.2014.03.009

Redondo, J., Fraga, I., Padrón, I., & Comesaña, M. (2007). The Spanish adaptation of ANEW (affective norms for English words). *Behavior Research Methods*, *39*(3), 600–605. https://doi.org/10.3758/BF03193031

Rickard, N. S. (2004). Intense emotional responses to music: a test of the physiological arousal hypothesis. *Psychology of Music*, *32*(4), 371–388. https://doi.org/10.1177/0305735604046096

Rottenberg, J., & Gotlib, I. (2004). Socioemotional Functioning in Depression. In *Mood Disorders: A Handbook of Science and Practice* (pp. 61–77).

Rottenberg, J., Gross, J. J., & Gotlib, I. H. (2005). Emotion context insensitivity in major depressive disorder. *Journal of Abnormal Psychology*, *114*(4), 627–639. https://doi.org/10.1037/0021-843X.114.4.627

Rottenberg, J., & Hindash, A. C. (2015). Emerging evidence for emotion context insensitivity in depression. *Current Opinion in Psychology*, *4*, 1–5. https://doi.org/10.1016/j.copsyc.2014.12.025

Rush, J., Trivedi, M. H., Ibrahim, H. M., Carmody, T. J., Arnow, B., Klein, D. N., … Keller, M. B. (2003). The 16-Item Quick Inventory of Depressive Symptomatology. *Depression*, *54*(5), 573–583. https://doi.org/10.1016/S0006-3223(03)01866-8

Russell, C. J., & Bobko, P. (1992). Moderated Regression Analysis and Likert Scales Too Coarse for Comfort. *Journal of Applied Psychology*, *77*(3), 336–342.

Russell, J. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*. Retrieved from http://psycnet.apa.org/psycinfo/1981-25062-001

Sandell, K., & Bornäs, H. (2017). Functioning Numbness Instead of Feelings as a Direction: Young Adults' Experiences of Antidepressant Use. *Sociology*, *51*(3), 543–558. https://doi.org/10.1177/0038038515591947

Santiago-Rivera, A. L., & Altarriba, J. (2002). The Role of Language in Therapy With the Spanish-English Bilingual Client. *Professional Psychology: Research and Practice*, *33*(1), 30–38. https://doi.org/10.1037/0735-7028.33.1.30

Schutte, N. S., Malouff, J. M., Bobik, C., Coston, T. D., Greeson, C., Jedlicka, C., … Wendorf, G. (2001). Emotional intelligence and interpersonal relations. *Journal of Social Psychology*, *141*(4), 523–536. https://doi.org/10.1080/00224540109600569

Scott, G., O'Donnell, P., Leuthold, H., & Sereno, S. (2009). Early emotion word processing: evidence from event-related potentials. *Biological Psychology*, *80*(1), 95–104. https://doi.org/S0301-0511(08)00073-2 [pii]\r10.1016/j.biopsycho.2008.03.010

Scott, G., O'Donnell, P., & Sereno, S. (2012). Emotion words affect eye fixations during reading. *Journal of Experimental Psychology: Learning Memory and Cognition*, *38*(3), 783–792. https://doi.org/10.1037/a0027209

Segalowitz, N., Trofimovich, P., Gatbonton, E., & Sokolovskaya, A. (2008). Feeling affect in a second language: The role of word recognition automaticity. *The Mental Lexicon*, *3*(1), 47–71. https://doi.org/10.1075/ml.3.1.05seg

Sharot, T. (2011). The optimism bias. *Current Biology*, *21*(23), R941–R945. https://doi.org/10.1016/j.cub.2011.10.030

Sheikh, N. A., & Titone, D. A. (2013). Sensorimotor and linguistic information attenuate emotional word processing benefits: An eye-movement study. *Emotion, 13*(6), 1107–1121. https://doi.org/10.1037/a0032417

Sheikh, N., & Titone, D. (2015). The embodiment of emotional words in a second language: An eye-movement study. *Cognition & Emotion,* (April), 1–13. https://doi.org/10.1080/02699931.2015.1018144

Shields, M., Gorber, S. C., Janssen, I., & Tremblay, M. S. (2011). Bias in self-reported estimates of obesity in Canadian health surveys: An update on correction equations for adults. *Health Reports*, *22*(3).

Sigmon, S. T., & Nelson-Gray, R. O. (1992). Sensitivity to Aversive Events in Depression: Antecedent, Concomitant, or Consequent? *Journal of Psychopathology and Behavioral Assessment*, *14*(3), 225–246.

Skehan, P. (1991). Individual Differences in Second Language Learning. *Studies in Second Language Acquisition*, *13*(02), 275. https://doi.org/10.1017/S0272263100009979

Sloan, D. M., Strauss, M. E., Quirk, S. W., & Sajatovic, M. (1997). Subjective and expressive emotional responses in depression. *Journal of Affective Disorders*, *46*, 135–141. https://doi.org/https://doi.org/10.1016/S0165-0327(97)00097-9

Sloan, D. M., Strauss, M. E., & Wisner, K. L. (2001). Diminished Response to Pleasant Stimuli by Depressed Women. *Journal of Abnormal Psychology*, *110*(3), 488–493. https://doi.org/I0.1037//0021-843X.110.3.488

Smith, D. P., Hillman, C. H., & Duley, A. R. (2005). Influences of Age on Emotional Reactivity During Picture Processing. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *60*(1), P49–P56. https://doi.org/10.1093/geronb/60.1.P49

Soares, A. P., Comesaña, M., Pinheiro, A. P., Simões, A., & Frade, S. (2012). Behavior and Research Methods Corresponding author : *Behavior Research Methods*, *44*(1), 256–269. https://doi.org/10.3758/s13428-011-0131-7

Stadthagen-González, H., Ferré, P., Pérez-Sánchez, M. A., Imbault, C., & Hinojosa, J. A. (2017). Norms for 10,491 Spanish words for five discrete emotions: Happiness, disgust, anger, fear, and sadness. *Behavior Research Methods*. https://doi.org/10.3758/s13428-017-0962-y

Stadthagen-Gonzalez, H., Imbault, C., Pérez Sánchez, M. A., & Brysbaert, M. (2017). Norms of valence and arousal for 14,031 Spanish words. *Behavior Research Methods*, 1–45. https://doi.org/10.3758/s13428-015-0700-2

Statistics Canada. (2013). *National Household Survey*. *Immigration and Ethnocultural Diversity in Canada* (Vol. 99-010-X20). https://doi.org/99-010-X2011001

Sutton, T. M., Altarriba, J., Gianico, J. L., & Basnight-Brown, D. M. (2007). The automatic access of emotion: Emotional Stroop effects in Spanish-English bilingual speakers. *Cognition and Emotion*, *21*(5), 1077–1090. https://doi.org/10.1080/02699930601054133

Thoma, P., Norra, C., Juckel, G., Suchan, B., & Bellebaum, C. (2015). Performance monitoring and empathy during active and observational learning in patients with major depression. *Biological Psychology*, *109*, 222–231. https://doi.org/10.1016/j.biopsycho.2015.06.002

Tsai, J. L., Knutson, B., & Fung, H. H. (2006). Cultural variation in affect valuation. *Journal of Personality and Social Psychology*, *90*(2), 288–307. https://doi.org/10.1037/0022-3514.90.2.288

Velez-Uribe, I., & Rosselli, M. (2019). The auditory and visual appraisal of emotion-related words in Spanish–English bilinguals. *Bilingualism: Language and Cognition*, *22*(1), 30–46. https://doi.org/10.1017/S1366728917000517

Vinson, D., Ponari, M., & Vigliocco, G. (2014). How does emotional content affect lexical processing? *Cognition & Emotion*, *28*(4), 737–746. https://doi.org/10.1080/02699931.2013.851068

Võ, M. L. H., Jacobs, A. M., Kuchinke, L., Hofmann, M., Conrad, M., Schacht, A., & Hutzler, F. (2008). The coupling of emotion and cognition in the eye: Introducing the pupil old/new

effect. *Psychophysiology*, *45*(1), 130–140. https://doi.org/10.1111/j.1469-8986.2007.00606.x

Wang, Y. guang, Wang, Y. qiang, Chen, S. lin, Zhu, C. yan, & Wang, K. (2008). Theory of mind disability in major depression with or without psychotic symptoms: A componential view. *Psychiatry Research*, *161*(2), 153–161. https://doi.org/10.1016/j.psychres.2007.07.018

Waris, O., Soveri, A., Ahti, M., Hoffing, R. C., Ventus, D., Jaeggi, S. M., … Laine, M. (2017). A latent factor analysis of working memory measures using large-scale data. *Frontiers in Psychology*, *8*(JUN), 1–14. https://doi.org/10.3389/fpsyg.2017.01062

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191–1207. https://doi.org/10.3758/s13428-012-0314-x

Warriner, A. B., Shore, D. I., Schmidt, L. A., Imbault, C. L., & Kuperman, V. (2017). Sliding into happiness: A new tool for measuring affective responses to words. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *71*(1), 71–88. https://doi.org/10.1037/cep0000112

Watson, D., Clark, L. A., & Carey, G. (1988). Positive and negative affectivity and their relation to anxiety and depressive disorders. *Journal of Abnormal Psychology*, *97*(3), 346–353. https://doi.org/10.1037/0021-843X.97.3.346

Weisleder, A., & Fernald, A. (2013). Talking to Children Matters: Early Language Experience Strengthens Processing and Builds Vocabulary. *Psychological Science*, *24*(11), 2143–2152. https://doi.org/10.1177/0956797613488145

Westbury, C., Keith, J., Briesemeister, B. B., Hofmann, M. J., & Jacobs, A. M. (2015). Avoid violence, rioting, and outrage; approach celebration, delight, and strength: Using large text corpora to compute valence, arousal, and the basic emotions. *The Quarterly Journal of Experimental Psychology*, *0*(0), 1–24. https://doi.org/10.1080/17470218.2014.970204

Wickham, H., & Chang, W. (2016). Create elegant data visualisations using the grammar of graphics. https://doi.org/10.1093/bioinformatics/btr406

Williams, J. M., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and

psychopathology. *Psychological Bulletin*, *120*(1), 3–24. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8711015

Winskel, H. (2013). The emotional Stroop task and emotionality rating of negative and neutral words in late Thai-English bilinguals. *International Journal of Psychology*. Taylor & Francis. https://doi.org/10.1080/00207594.2013.793800

Wolkenstein, L., Schönenberg, M., Schirm, E., & Hautzinger, M. (2011). I can see what you feel, but i can’t deal with it: Impaired theory of mind in depression. *Journal of Affective Disorders*, *132*(1–2), 104–111. https://doi.org/10.1016/j.jad.2011.02.010

Wolter, B. (2001). COMPARING THE L1 AND L2 MENTAL LEXICON A Depth of Individual Word Knowledge Model. *Studies in Second Language Acquisition*, *23*(1), S0272263101001024. https://doi.org/10.1017/S0272263101001024

Zuroff, D., Koestner, R., Moskowitz, D. S., McBride, C., Marshall, M., & Bagby, M. (2007). Autonomous motivation for therapy: A new common factor in brief treatments for depression. *Psychotherapy Research*, *17*(2), 137–148. https://doi.org/10.1080/10503300600919380

Chapter 2

Abstract

The topic of non-native language processing has been of steady interest over past decades. Yet, conclusions about the emotional responses in L2 have been highly variable. We conducted a large-scale rating study to explicitly measure how non-native readers of English respond to the valence and arousal of 2,628 English words. We investigated the effect of a rater's L2 proficiency on how native-like their ratings are, and the way the semantic category of the word affects how L2 readers experience and rate that word. Non-native speakers with a longer time lived in Canada showed emotional responses that were more similar to those by L1 speakers of English. Similarly, L2 speakers with greater self-reported proficiency in English showed similar emotional responses to native speakers than L2 speakers with weaker self-reported English proficiency. Additionally, we saw that the valence differences between L1 and L2 raters were exacerbated in words that L2 raters typically do not use in English. These findings highlight the importance of behavioural ecology in language learning.

1. Introduction

An important question within the language sciences concerns how production, comprehension, learning, and memory storage of language can be modulated by extra-linguistic factors such as emotion. An equally intriguing question is whether emotion conveyed in one's first language (L1) has a different influence on cognitive and linguistic processing than in one's second language (L2). The relationship between L1 processing and emotion has been in the focus of psycholinguistics for the past decade, and extensive literature shows that emotional words are often named and recognized faster, and are more easily remembered, than neutral words (Altarriba & Bauer, 2004; Bayer et al., 2011; Briesemeister et al., 2011; Kousta et al., 2009; Kuperman et al., 2014; Long & Titone, 2007; Mackay et al., 2004; Recio et al., 2014; Scott et al., 2009; N. A. Sheikh & Titone, 2013; Vinson et al., 2014). Effects of word emotionality have also been found physiologically and in brain imaging (Citron, Weekes, & Ferstl, 2013). For instance, Recio et al. (2014) found increased early posterior negativity for positive and arousing words than neutral or negative words, suggesting quicker processing. Additionally, Bayer et al. (2011) found that highly arousing words led to greater pupillary constriction than low arousing words.

The magnitude and even direction of emotionality effects on word processing demonstrably depends on several factors, which include the nature of the task. For instance, tasks where word emotionality is irrelevant, emotionality is distracting and results in slower processing (e.g., emotional Stroop tasks; Eilola, Havelka, & Sharma, 2007; Sutton, Altarriba, Gianico, & Basnight-Brown, 2007). As well, many studies obtain consistent advantage in word recognition speed for both positively and negatively valenced words (Kousta et al., 2009; Kuchinke, Võ, Hofmann, & Jacobs, 2007; Méndez-Bértolo, Pozo, & Hinojosa, 2011; Scott et al., 2009; Scott, O'Donnell, & Sereno, 2012), whereas others find that only positivity comes with a processing boost (Kuperman et al., 2014; Vinson et al., 2014). Yet, there appears to be a consensus that word emotionality has a noticeable impact on language processing.

Differences in the emotional response to words in L1 and L2 have also been widely reported across language and tasks, with contested and variable outcomes. A general observation is that when a person hears or reads the word "death" in their L1, English, some psycholinguistic work suggests that they experience strong emotions and memories associated with the word (reviewed in Itzhak, Vingron, Baum, & Titone (2017). However, when the same person hears or reads the translation equivalent in their L2, French, (i.e., “mort”), the same strong emotions may be neutralized despite the person’s high familiarity with the word. Even highly proficient bilinguals report that they feel more emotional and natural in their first language than their second language, regardless of their L2 proficiency (J. M. Dewaele & Nakano, 2013; Santiago-Rivera & Altarriba, 2002). Yet vast individual differences between L2 speakers, and variability in tasks and materials, make existing findings less than consistent (Ayçiçeği-Dinn & Caldwell-Harris, 2009; Brase & Mani, 2017; Caldwell-Harris, 2015; De Grauwe, Willems, Rueschemeyer, Lemhöfer, & Schriefers, 2014; Foroni, 2015; Ponari, Rodriguez-Cuadrado, Vinson, et al., 2014). Accordingly, many studies have shown that non-native speakers do not experience this same processing boost of emotional stimuli like native speakers. For example, Anooshian & Hertel (1994) found that Spanish-English bilinguals recalled more emotional stimuli in their first language, but not in their second language. Furthermore, Caldwell-Harris, Tong, Lung, & Poo (2011) found that Chinese-English bilinguals rated Chinese (L1) phrases as more emotional than English (L2) phrases, but found similar skin conductance responses to both languages. Similarly, sequential German-French bilinguals only experience affective priming in their L1, and only bilinguals with high levels of L2 immersion experienced affective priming in their L2 (Degner, Doycheva, & Wentura, 2012). Further evidence of emotional neutralization in L2 comes from studies on reprimands and taboo phrases, where L2 (English) phrases elicited smaller skin conductance responses (SCRs) than L1 (Turkish) phrases (Harris et al., 2003).

Unlike the studies above, other research has shown no processing difference between L1 and L2 processing of emotional and neutral words. In contrast to Anooshian and Hertel (1994), Ferré, García, Fraga, Sánchez-Casas, & Molero (2010) did not find a difference in the number of emotional words recalled in L1 or L2 of early Spanish-Catalan (and Catalan-Spanish) bilinguals and late Spanish-English bilinguals, suggesting intact emotional processing regardless of when the experiencer learned the language (for similar results in Turkish-English bilinguals, see Ayçiçeği-Dinn & Caldwell-Harris, 2009). As well, Ponari et al. (2014) found that highly proficient L2 speakers of English showed similar emotional facilitation in a lexical decision task as L1 speakers. In an emotional Stroop task, Spanish-English bilinguals experienced interference in both languages, with the effect being larger in their second language, English (Sutton et al., 2007). Sutton et al. (2007) propose that this is due to the frequency of language use, with most participants using English more than they do Spanish (L1). A similar finding was shown in Eilola et al. (2007).

These inconsistencies in the L2 literature may arise from differences in the level of dominance and proficiency in L1 and L2, as well as the type of emotional words used in the tasks. Kazanas & Altarriba (2016) found processing differences between emotion (e.g., happiness) and emotion-laden (e.g., vacation) in one's dominant language, but not necessarily in their first language. Harris (2004) found similar results in early Spanish-English bilinguals. Early bilinguals (who used English, their non-native language, dominantly) showed no difference in SCRs between English and Spanish. With respect to word type, an eye-tracking study by Sheikh & Titone (2015) found that negative words had less of an impact on L2 reading times, but the impact of positive words was spared. Moreover, Sheikh and Titone (2015) found that L2 reductions in negative word processing were modulated by both word frequency and abstractness in both L1 and L2 processing. That is, for high frequency abstract words, Sheikh and Titone found reading facilitation for both positive and negative words; however, only positive words were facilitated for L2 readers. In contrast, for low frequency words, L1 readers were faster for all words, and L2 readers showed facilitation of positive words, irrespective of abstractness.

Several theories offer insights into how people's real-world usage with emotionally charged words can impact the way that they are processed in the L1 and L2. Specifically, L1 and L2 language users have differential experience with positive vs. negative words in their linguistic environments: as one possibility, bilinguals may prefer to swear or talk about feelings in L1, and conduct professional interaction or receive education in L2. These differences in the so-called "behavioral ecology" interact with other word attributes (e.g., frequency, abstract/concreteness). With respect to reading in particular, these experiential differences in L1 vs. L2 processing may be accounted for by two alternative theories. The first theory, Perfetti's (2007) Lexical Quality Hypothesis, posits that one's level of proficiency and word knowledge can impact natural reading. Accordingly, differences between L1 and L2 readers emerge when non-native speakers do not know some words, and so they choose to rate these words as less emotional. This generates a prediction that less frequent words should result in a larger neutralization of emotional responses (and hence, ratings) for non-native speakers of English (less frequent words are less likely to be encountered by a non-native speaker). Additionally, less proficient L2 speakers of English should also demonstrate more neutralization of affect (i.e., less extreme ratings) than more proficient L2 speakers of English. This prediction finds support in some studies showing that more proficient non-native speakers behaved like native speakers in emotional Stroop tasks (Eilola et al., 2007; Sutton et al., 2007). However, it conflicts with Degner et al. (2012 and Segalowitz, Trofimovich, Gatbonton, & Sokolovskaya (2008), who did not find any interaction between L2 neutralization and proficiency.

The second theory, based on Barsalou's (1999) Embodied Theoretical Approach to Language, Dual Coding Theory (Clark & Paivio, 1991) and recent work by Kousta, Vigliocco, Vinson, Andrews, & Del Campo (2011) suggests that learning the emotional connotation of a word does not always come with learning the definition (denotation) of that word, but instead is acquired by our interaction with those words in real-life environments, with its affective and sensorimotor aspects. For L2 readers to acquire the emotional connotations of foreign-language words, they need to be exposed to the ecologically valid, affective and sensorimotor experiences that come with that word. Simply knowing the denotation of a word (say, through book reading or classroom education) is not sufficient. The link between the embodied experience with a word and its meaning acquisition is made explicit in Dual Coding Theory (DCT, Clark & Paivio, 1991), among others. DCT proposes that abstract concepts are only stored as verbal representations, but concrete concepts are stored as both verbal and imaginal representations. Thus, concrete words are more readily available than abstract words because they are represented by two systems, whereas abstract words are only represented by one. This theory suggests that L2 speakers of a language should have intact emotion processing for concrete words (which they have experienced and have two levels of representation) but not abstract words, see Sheikh and Titone (2015). However, Kousta et al. (2011) propose that DCT cannot fully account for the processing of abstract concepts. They state that abstract concepts are often more emotional than concrete concepts, and that this emotion plays a role in how abstract words are processed. They suggest that sensorimotor information is most important for the representation of concrete words, whereas affective information is most important for abstract words.

Harris, Gleason, & Ayçiçeği (2006) argue that the context in which a word is learned has a strong effect on how emotional that word is to the experiencer. They proposed the Emotional Contexts of Learning Theory, which theorizes that it’s not the order of language learning that matters for emotional processing, but rather how emotional the learning context is. For most bilinguals, their first language is more emotional because it is learned in an emotional environment (such as with family). On the other hand, a second language can have variable levels of emotionality in its learning context. Some bilinguals may learn their second language in an equally emotional environment to their L1, while others may learn an L2 a less emotional (or personal) environment such as school or work.

In sum, the family of embodied accounts predicts that the environment a word is learned in, which we can proxy by using the semantic category of a word (e.g., homework as a work-related word vs. butt as a swear word; see, Pennebaker, Booth, & Francis, 2007), should affect the valence rating of that word for L2 readers. If a word falls into a semantic category that aligns with how L2 readers use English (e.g., at school or work, but potentially not at home with their family and friends), then the word will have as much emotionality as with L1 readers. If the word is not used by L2 readers, even if they know the meaning of the word, it will have less emotional impact on these individuals. In line with this theory, how a non-native speaker learned English should interact with how native-like their L2 emotion is. Specifically, after controlling for proficiency, non-native speakers who have lived in an English-speaking country for more years should have more native-like English emotion than those who recently moved to an English-speaking country.

The present study

Given the mixed results of empirical studies and alternative theoretical predictions, our goal is to fill a gap in the literature by conducting a large-scale data collection of affective ratings to English words from L2 readers of English. Among other needs, this study addresses a common practice of using affective norms collected from L1 speakers of a language to assess emotional and cognitive processing of that language in its L2 speakers.

To overcome limitations described above, the present study uses a large stimulus list to tease apart which linguistic factors affect the emotional processing of a word in L2, and more participants than any other study mentioned to understand which demographic variables affect how a speaker processes emotion, although all participants are undergraduate students with similar age and educational backgrounds. Additionally, we explicitly ask participants to rate the emotion of a word, rather than having them interact with emotional words in an unrelated task, to pinpoint the emotional differences between words. Thus, we specifically examine the emotional connotation of English words by directly asking participants to rate the valence and arousal of each word on a 9-point ratings scale (see Warriner, Kuperman, & Brysbaert, 2013). While large datasets exist for L1 affective norms in a variety of languages (Bonin, Aubert, Malardier, & Niedenthal, 2016; Soares, Comesaña, Pinheiro, Simões, & Frade, 2012; Stadthagen-Gonzalez, Imbault, Pérez Sánchez, & Brysbaert, 2017; Warriner et al., 2013, among many others), the data amassed in this study is the largest set of valence and arousal ratings given by non-native speakers of English.

The present study is not the first study where L1 and L2 affective ratings are compared. Velez-Uribe & Rosselli (2019) collected affective ratings in English and Spanish from 101 Spanish-English bilinguals for 120 words, consisting of 40 positive, 40 negative, and 40 taboo/swear words. They found that these bilinguals rated positive and negative words as more emotional in their second language (English), but rated taboo/swear words as more emotional in their first language (Spanish). Winskel (2013) collected affective ratings to Thai and English words from 57 Thai-English bilinguals. These participants rated the valence of 20 neutral and 20 negative words (comprised of negative emotion and negative emotional words). Language had no effect on the bilinguals' valence rating of the word. Both of these studies have small sample sizes and relatively homogeneous linguistic makeup, but provide interesting insight into the difference in explicit affective ratings between a bilingual's L1 and L2.

2. Materials and methods

Participants

Valence

420 undergraduate students, enrolled in linguistic courses at McMaster University, participated in the collection of valence norms for course credit. Of these, 14 participants identified their first language as English and were excluded from further analyses. We removed an additional 9 participants who provided "I don't know this word" responses to more than 50% of stimuli, as well as 137 participants who provided the same numeric response to 50% or more of stimuli.

Among the remaining 260 participants, 203 participants were female and 234 were right-handers (5 participants did not provide handedness information). The mean age of participants was 20.13 (sd = 3.67), and 94% of participants were between the ages of 18 and 24. The distribution of native languages (L1s) in the cohort was as follows: 99 (38%) participants reported Mandarin, Cantonese or an unspecified Chinese dialect as their L1; 25 (10%) reported Arabic; 20 (8%) Urdu; 19 (7%) Korean, and the remaining participants reported other single or multiple L1s.

See Table 1 in the Appendix for further demographic variables relating to participants who rated word valence.

Arousal

644 undergraduate students, enrolled in linguistic courses at McMaster University, participated in the collection of arousal norms for course credit. Of these, 24 participants identified their first language as English and were excluded from further analyses. We removed an additional 134 participants who provided “I don’t know” responses to more than 50% of stimuli, as well as 195 participants who provided the same numeric response to 50% or more of stimuli.

Among the remaining 291 participants, 216 participants were female and 258 were right-handers (7 participants did not provide handedness information). The mean age of participants was 20.25 (sd = 3.6), and 92% of participants were between the ages of 18 and 24. The distribution of native languages (L1s) in the cohort was as follows: 129 (44%) participants reported Mandarin, Cantonese or an unspecified Chinese dialect as their L1; 26 (9%) reported Arabic; 17 (6%) Urdu; 18 (6%) Korean, and the remaining participants reported other single or multiple L1s .

Materials and Design

The words included in our stimulus set were selected from the set of about 14,000 English words for which Warriner, Kuperman, and Brysbaert (2013) made available of affective norms collected amongst L1 speakers. In this study, valence ratings were collected for 2688 (2692 for arousal) unique words on a 9-point scale (for a definition of the scale, see below). The stimuli for the present study were distributed over 8 lists (one per semester) containing between 325 to 360 words each. The words were presented in a random order, and were randomly selected from Warriner et al.’s ratings to uniformly represent the range of valence or arousal. Specifically, we binned all words in Warriner et al. into 25 bins (5 bins of valence by 5 bins of arousal) and randomly drew 13 words from each bin without replacement, totaling approximately 325 words per session. Some bins depleted in subsequent lists. Once a bin was depleted, we randomly drew 15 words from each of the remaining bins without replacement. Six lists shared 10 anchor words at the start of each session, while 2 lists shared no words with any other list.

Procedure

Participants were recruited via the departmental online subject pool over 8 semesters from 2012-2017. Since a different (single) stimulus list was offered in each semester, participants were allowed to complete more than one experimental list in different semesters, and some participants completed the task for both valence and arousal. In the first two semesters, each word was simultaneously rated both for valence and arousal; in following sessions, words were separately rated for valence and arousal. The instructions to all stimuli lists were as follows, with wording for arousal in square brackets:

> *You are going to be asked to rate how happy [excited] or sad [calm] a series of words make you feel. There will be approximately 360 words. The scale ranges from 1 (happy [excited]) to 9 (unhappy [calm]). At one extreme of this scale, you are happy, pleased, satisfied, contented, hopeful [stimulated, excited, frenzied, jittery, wide-awake, or aroused]. When you feel completely happy [aroused] you should indicate this by choosing rating 1. The other end of this scale is when you feel completely unhappy, annoyed, unsatisfied, melancholic, despaired, or bored [relaxed, calm, sluggish, dull, sleepy, or unaroused]. You can indicate feeling completely unhappy [calm] by selecting 9. The numbers also allow you to describe intermediate feelings of pleasure [calmness/arousal], by selecting any of the other feelings. If you feel completely neutral, neither happy nor sad [not excited nor at all calm], select the middle of the scale [rating 5]. If you do not know a word well enough to rate it, please check the button for "Unknown Word".*

Prior to completing valence or arousal ratings, participants were asked to supply basic demographic information (age, gender, handedness). They also completed an extensive bilingualism questionnaire (LEAP-Q), which queried one’s language education and immigration background, one's proficiency in their first language, English and other non-native languages. They also indicated the situations and environments of each language use (Marian, Blumenfeld, & Kaushanskaya, 2007). Completion of the entire task did not exceed 45 minutes.

2. Results

Data Trimming

A total of 91658 valence (104058 for arousal) ratings were collected; the number was reduced to 88224 (97315 for arousal) after we removed ratings that were outside the 1-9 scale (including the 3125 responses signifying an unknown word; 3364 for arousal). We further removed all words that received 12 or fewer ratings (15.03% of valence ratings, 0% of arousal ratings). We reversed the valence [arousal] scale from the one presented to participants; 1 represents unhappy [calm] and 9 represents happy [excited] for more intuitive interpretation.

The resulting pool contained 83590 valence (93451 arousal) ratings to 2284 English words submitted by 260 (289 for arousal) L2 speakers of English. Experimental sessions varied in the number of recruited participants. Each word received between 13 and 228 ratings with a median of 27, and a mean of 36.6 (*SD* = 26.2). This disparity was due to earlier sessions using anchor words at the start of the study (resulting in a large number of ratings), and later sessions using no anchors. We provide a spreadsheet reporting each word, its average valence and arousal rating, the standard deviation of both ratings, and the number of ratings per word (outlined in the Availability section). We further supplement this information by the average, standard deviation and number of valence and arousal ratings given to those words by all L1 speakers of English in Warriner et al.'s (2013) study, as well as by the subgroup of younger (less than 30) participants in that study: this subgroup is age-compatible with our present cohort. Finally, we supply the word's frequency of occurrence in the 51-million token corpus of subtitles to English films and media (Brysbaert & New, 2009)..

We estimated the split-half reliability of L2 valence ratings using a bootstrapping procedure. At each of 1000 iterations, we randomly selected two sets of 6 ratings to each word without replacement (half of the minimum number of ratings to a word: 13), and calculated the mean rating in each group. This provided us with two sets of 2627 mean ratings: the correlation coefficient between each two sets is a point-wise estimate of split-half reliability, and the distribution of coefficients over 1000 iterations enabled an estimation of the confidence interval. L2 valence ratings showed a relatively high split-half reliability: the average correlation was r = 0.74 [CI 95% 0.72–0.75]. This reliability was lower than one reported in Warriner et al. (2013), r = 0.91. This is likely to greater heterogeneity in the level of English proficiency, as well as linguistic, cultural and social backgrounds amongst our L2 raters, compared to the native English speakers from the US that were recruited in Warriner et al. (2013).

We performed the split-half reliability of L2 arousal ratings using the same method as described above. L2 arousal ratings showed a weak split-half reliability: the average correlation was r = 0.34 [CI 95% 0.31-0.37]. In line with Warriner et al. (2013), split-half reliability of arousal ratings is substantially lower than the reliability of valence ratings (r = 0.689 in Warriner et al. 2013)). Due to the low reliability of the arousal ratings, we will not report any analyses henceforth.

Comparison of First and Second Language Ratings

This section conducts comparisons of overall distributions of valence ratings given to 2627 English words by L2 and L1 speakers of English, as documented in the present study and in Warriner et al. (2013), respectively. We then consider valence ratings of both groups to specific semantic categories that are hypothesized to bring forward differences in behavioral ecology of native and non-native speakers of a language: e.g., work-related versus swear words. Finally, we consider the moderating influence of one's L1, one's proficiency in L1 and L2 and other pertinent factors. Given the age distribution of the L2 participants, we chose to compare their judgments to those of the *young* (under 30) subset of L1 raters from Warriner et al.'s (2013) data set: in what follows, all references to L1 ratings relate to this subset. We note that comparisons with ratings based on the entire L1 population in Warriner et al. yielded nearly identical results.

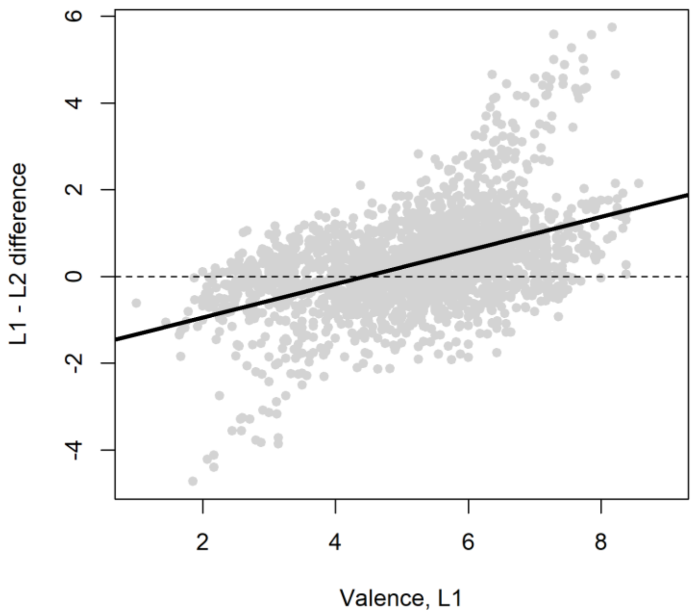

Figure 1 Scatterplot of the difference between L1 and L2 valence ratings plotted against L1 ratings, with a lowess trend line (solid).

To zoom in on differences in valence judgments, we calculated the difference between the mean L1 and L2 valence ratings (L1 - L2). A negative difference signifies a word that received a more positive response from non-native speakers of English; a positive difference means that the L1 valence rating was more positive (higher) than the L2 rating. Figure 1 plots the difference score against L1 valence ratings. The dotted line in this figure marks the hypothetical position of identical ratings from L1 and L2 speakers of English. The best-fit regression line demonstrated a linear effect of L1 valence on difference scores ($y = -0.91 + 0.21x$, $R^2 = 0.14\%$. Most words in the middle of the valence range are close to this equality line, yet, the further a word is from being neutral, the more L2 valence ratings deviate from L1 ratings.. L2 responses were consistently more moderate than L1 responses at the ends of the scale (less positive for positive words and less negative for negative ones). The more extreme the L1 judgments were, the more attenuated the L2 response was. Thus, for words with L1 ratings below 2 (very negative) or above 8 (very positive) on a 9-point scale, L2 speakers provided judgments that were about 1 point higher (for negative) or 1 point lower (for positive) respectively. Apparent exceptions to this pattern are words at the very extremes of the valence distribution: the most negative words in our stimulus list (e.g., *betrayal*, *bankrupt*, L1 valence 1.00 and 1.44) and the most positive ones (e.g., *happy, relaxing*, L1 valence 8.38 and 8.38) elicited L2 valence judgments that were closer to L1 valence judgments than words slightly further from the extremes. The correlation between L1 valence judgements and the L1-L2 difference was 0.48, $df = 2282$, $p < 0.001$.

L2 Proficiency and word frequency as moderators of emotionality

A straightforward prediction for the proficiency-driven discrepancy in L1 and L2 emotionality is that the discrepancy would be stronger in lower-frequency words and/or in bilinguals with a lower level of L2 proficiency. We tested this hypothesis by examining how the difference between L1 and L2 ratings of valence is influenced by a word's frequency of occurrence in language, and estimates of English proficiency of our raters.

From the pool of valid responses to 2008 words, we selected 73576 ratings given by 223 participants who specified all aspects of their English proficiency in our bilingualism questionnaire. We selected the participant's subjective estimate of English reading mastery as an index of their English proficiency. We fit a linear mixed-effects multiple regression model to the difference between L1 and L2 ratings as a dependent variable. The model contained a critical interaction of two fixed effects: L1 rating of valence and English reading proficiency and a main effect of word frequency. We centered the estimates of L1 valence and log transformed word frequency for improved interpretability of the model. Random effects included random intercepts by word and by participant. Library lmerTest (Kuznetsova, Brockhoff & Christensen, 2017) in the statistical software language R (R Core Team, 2018) was used in all regression analyses.

Subjective measures of English proficiency showed a statistically reliable interaction. The greater a participant's self-reported reading proficiency, the more native-like their ratings were (see Figure 2 for the interaction plot and Table 2 in the appendix for specification of the model). Furthermore, we found a significant effect of word frequency – more frequent words were rated more native-like than less frequent words. In sum, the attenuated magnitude of emotional responses in L2 speakers is observed in participants with low subjective English proficiency and less frequent words, but less so in highly proficiency L2 English speakers.

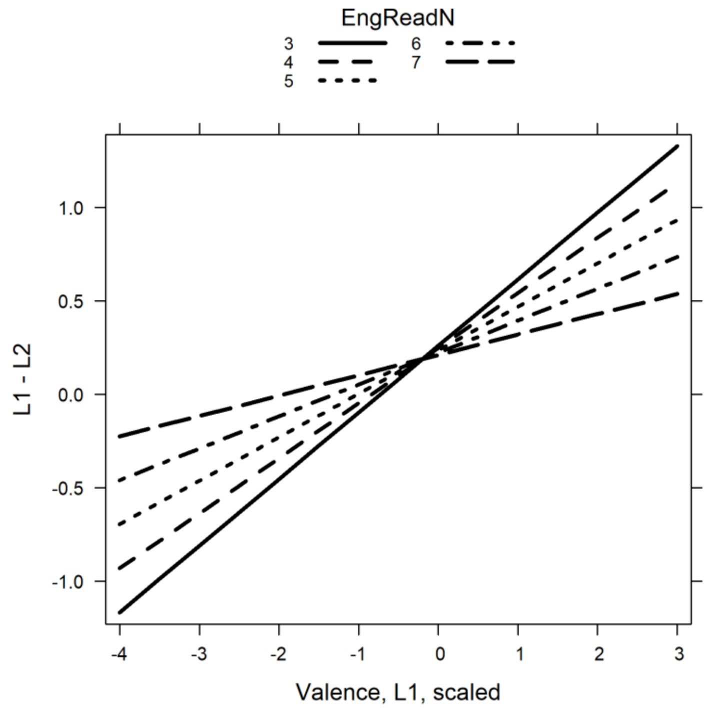

Figure 2 Effect of L1 valence ratings on the L1-L2 difference in valence ratings, broken down by 5 levels of subjective English reading proficiency.

Embodied emotion

We fit a separate linear mixed-effects multiple regression model to the difference between L1 and L2 ratings as a dependent variable and time spent in Canada as an independent variable. We centered the estimates of L1 valence and square-root transformed months in Canada for improved interpretability of the model. Random effects included random intercepts by word and by participant. Time lived in Canada, used as a proxy for the amount of English embodiment, demonstrated a reliable modulating effect on the relationship between L1 valence ratings and the L1-L2 difference, see Figure 3 for the interaction plot and Table 3 in the Appendix for specification of the model. Non-native speakers with a longer time lived in Canada showed emotional responses that were more similar to those by L1 speakers of English, see Figure 3.

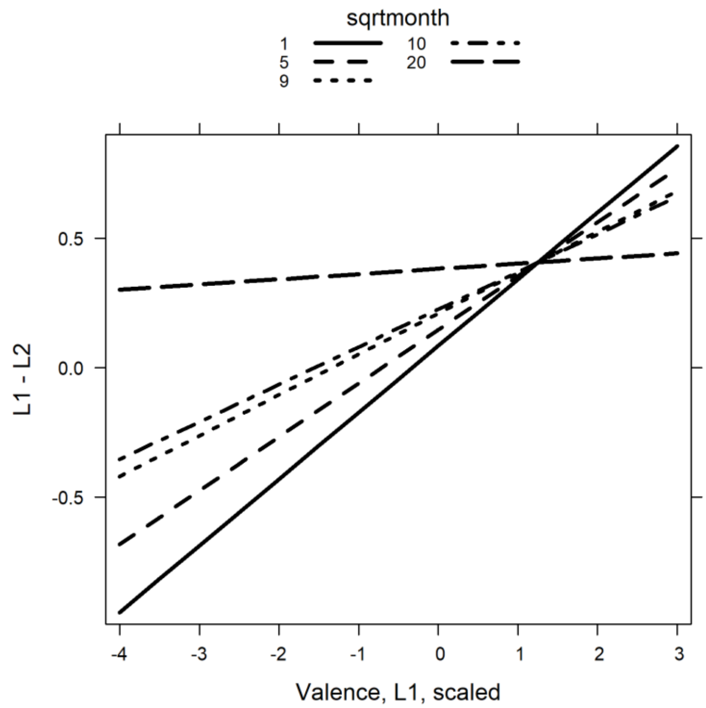

Figure 3 Effect of L1 valence ratings on the L1-L2 difference in valence ratings broken down by five levels of the time lived in Canada (in square-root transformed months).

We cannot combine both the proficiency account and the embodiment account into a single model. Subjective English proficiency and length of time in Canada are strongly correlated (0.54, df = 70347, $p < 0.001$). It is therefore infeasible to tease apart the effects of these two crucial predictors. It is likely that both English proficiency and English embodiment play overlapping and separate roles in the processing of English emotionality. Furthermore, our length of time in Canada estimate is a more accurate measure than subjective English proficiency rating. This could further lead to variability within our results and increase our inability to compare proficiency to embodiment.

To further investigate the effects of embodiment on emotionality, it is possible that a non-native speaker primarily uses their L1 rather than L2 to discuss topics related to swear or taboo words, but predominantly uses L2 rather than L1 to communicate in a professional setting. In this case, emotionality of home-related words in their L2 will deviate more from that in their L1, whereas their work-related lexicon might elicit similarly strong or even stronger emotional responses in their L2 than in their L1. This approach predicts a substantial variability in how similar L1 and L2 emotional responses to words are, depending on the semantic domains of the lexicon.

We tested this account by examining whether the discrepancy between L1 and L2 ratings varied by topic. To this end, we selected several topics from the LIWC, a dictionary of over 5 thousand words distributed over semantic categories (Pennebaker, Booth & Francis, 2007). For each topic we sought to establish whether the difference between L1 and L2 judgments of valence was stronger than that observed across the entire word set.

We cannot fully account for circumstances in which our participants tend to use English or their L1. For this reason, we selected for examination the topics that the prior literature found to elicit much weaker emotional responses in L2 speakers than in L1 speakers: e.g., swear, taboo and sex-related words, and words denoting positive and negative emotions. Additionally, given that our cohort consisted of bilingual students in a Canadian university, we assumed that topics related to occupation and achievement would correspond to the experiences in which our participants use English equally or more often than their L1. Furthermore, these professional contexts where our bilingual students use English are less conducive to using words we expect to be less emotional (swear, taboo and sex-related words). If a group of words are less emotional in our bilingual's L2, we expect a weak-to-no correlation between the difference in L1 and L2 ratings and the L2 rating of each word. If a group of words are more emotional in our bilingual's L2, we expect a stronger correlation between L1-L2 and L2 rating.

Data fully supported predictions of the embodied approach, and in agreement with prior literature demonstrating that swear and taboo words become less emotional with each subsequently learned language among multilinguals (Dewaele, 2004). 26 swear, taboo and sex-related words overlapped across the LIWC category "swear" and our word list. These words showed the strongest difference between L1 and L2 judgments. Figure 4a plots the L1 - L2 difference in valence ratings against L1 valence ratings for these words, with jitter between words added for legibility. The plot shows that taboo words judged as positive by L1 speakers (e.g., *cock*, *screw*, *butt*) were consistently judged as much less positive by L2 speakers; while very negative L1 judgments (e.g., *douche*, *faggot*, *prick*) rendered a much less negative response from non-native speakers of English. The correlation between the L1 valence ratings and the L1-L2 difference in ratings was 0.81(df = 24, $p < 0.001$). As indicated by the Fisher r-to-z transformation, this correlation was significantly stronger ($p < 0.001$) than the one observed in the entire dataset (0.48, df = 2282, $p < 0.001$).

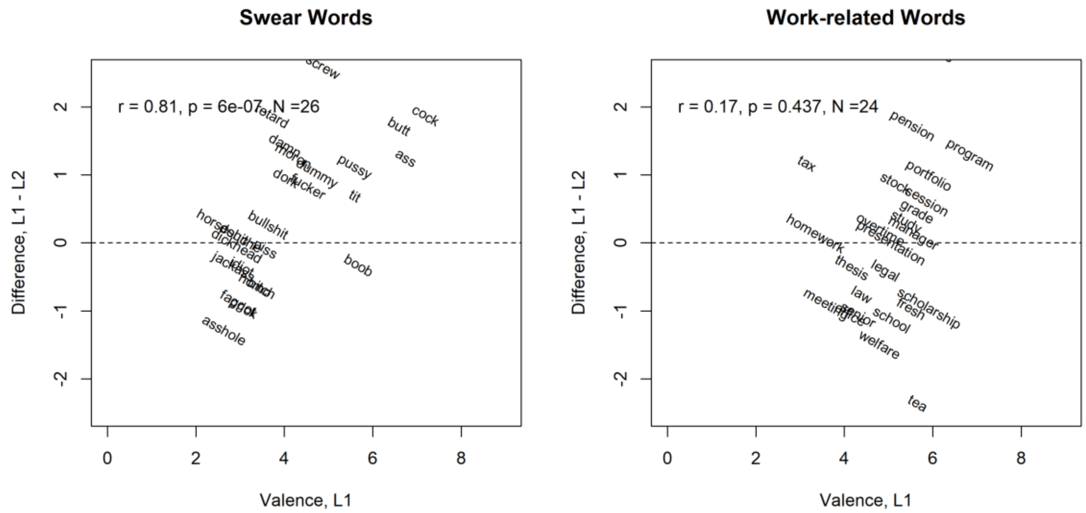

Figure 4 Difference between L1 and L2 valence ratings to a. swear, taboo and sex-related words and b. work-related words plotted against L1 valence ratings to those words. Jitter added for legibility.

A similar tendency was observed in the comparison of L1 and L2 ratings to words from the “negative emotion" and “positive emotion" categories in the LIWC dictionary, see Figures S1 and S2 in the Supplementary Materials. A set of 104 LIWC words associated with negative emotions showed a correlation of 0.49 (df = 116, $p < 0.001$) between the L1-L2 valence difference and L1 ratings, indicating again a significantly stronger (the Fisher r-to-z transformation $p < 0.001$) tendency for non-native speakers of English to elicit less extreme emotional responses to words with extremely unpleasant meanings. In a somewhat weaker way, non-native speakers showed a less positive response to 66 words associated with positive emotions, as compared to L1 speakers of English: the critical correlation was 0.3(df = 74, $p < 0.001$), which is significantly stronger than the baseline correlation observed in the entire dataset.

In agreement with our predictions about the preferred use of English in the professional life of our bilinguals, we found an unreliable correlation (0.17, df = 24, p = 0.437) between the L1-L2 difference in valence ratings and L1 valence ratings in 24 words from the LIWC “work” category, see Figure 4b. That is, non-native speakers were as emotionally engaged in responding to items related to work, job and occupation as native speakers of English were. This trend was not significantly different from the baseline correlation that we observed in the entire dataset (0.48, df = 2282, $p < 0.001$), thus we conclude that L2 speakers showed no emotional attenuation in the professional lexical domain.

It could be possible that one of the reasons we find a difference in swear and taboo words, but not work-related words, is that swear words have a larger range of available emotions, and work-related words have a limited-range of emotions available. We found the opposite case; our work-related words had a larger L1 valence range than swear words (work: range = 3.00-7.49, mean = 5.30, SD = 1.20; swear: range = 2.33-5.76, mean = 4.44, SD = 1.20). The differences in L1 and L2 ratings are not due to the limited range of work-related words, but seemingly the contexts in which these words arise. We speculated that the topic of achievements would also be one that our participants would discuss more often in English than in their L1. Figure S3 in the Supplementary Materials confirms this conjecture. There was no reliable correlation (0.28, df = 28, p = 0.46) between the L1-L2 valence difference and L1 valence ratings in the 31 words from the LIWC “achievement” category.

4. Discussion

We report the largest available dataset of English affective norms (valence and arousal) collected from non-native speakers of English, available in the Supplementary Materials. Overall, we found a general tendency for L2 speakers to have attenuated valence judgements compared to L1 speakers. With a possible exception of well-anchored extremely emotional words, L2 speakers appear to be less emotionally engaged when judging words in a foreign language. In contrast to many earlier studies that concentrated on only negative words in their comparison of L1 and L2 emotionality (but see Sheikh & Titone, 2015), we examined the differences across the entire valence range and observed that the pattern of attenuating emotionally extreme responses is present at both ends of the valence scale, but is more salient in the negative subrange of valence.

These findings are in contrast with two previously collected datasets comparing explicit affective ratings to words in a bilingual's L1 and L2 (Velez-Uribe & Rosselli, 2019; Winskel, 2013). Velez-Urive & Rosselli (2019) found that bilinguals rated positive and negative words as more emotional in English (L1), and found taboo words more emotional in Spanish (L2). Winskel (2013), on the other hand, found no difference between Thai-English bilinguals' ratings of neutral and negative English and Thai words. These differences can be accounted for by two reasons. Firstly, both of these studies had small, linguistically homogeneous samples, and secondly, both compared a small sample of words. Furthermore, both studies had their bilinguals rate both languages, which could affect how semantically linked the L1 and L2 translations of each word were in the study.

We proposed two opposing theories regarding these differences in emotional language between L1 and L2 speakers. The first theory, based off of Perfetti's (2007) Lexical Quality Hypothesis, proposed that the difference between L1 and L2 readers of English is simply due to a degraded mental representation of the word. Due to their less extensive and diverse experience with words in L2, bilinguals may be expected to demonstrate a less precise and entrenched knowledge of both the denotations and emotional connotations of those words than their L1 counterparts (Hellman, 2011; Wolter, 2001). This semantic knowledge would be particularly incomplete in situations that do not provide sufficient learning opportunities: either words only occurring rarely in language use, or bilinguals not having substantial exposure to L2, or both. In order to support this theory, our data would have to suggest that the word's frequency should play a vital role in the difference in L1 and L2 readers. Additionally, L2 proficiency should play a role in how L2 raters respond to the word - if the word is well represented in our L2 rater's mind (higher proficiency), more proficient L2 readers should rate words similarly to L1 raters.

Our data supported these claims - we found a significant effect of the frequency of a word impacted the way L2 raters responded to the emotion of that word. More frequent words resulted in more native-like affective responses. Furthermore, self-rated L2 reading proficiency significantly interacted with the difference between L1 and L2 raters – more proficient raters gave responses that resembled L1 raters.

The second hypothesis, centered around Barsalou's (1999) Embodied Theoretical Approach to Language, as well as Dual Coding Theory (Clark & Paivio, 1991) and recent work by Kousta et al. (2011) proposes that word representation is encoded not only by meaning, but also by our sensorimotor and emotional interactions with that word. That is, traces of episodic memory that encode a word's orthography, phonology and semantics during exposure to a word also encode the bodily and emotional states of the person during that exposure (Kousta et al., 2011; Clark & Paivio, 1991). This approach highlights the role of behavioural ecology, i.e., the differential use of one's native and non-native languages in a spectrum of communication settings. To support this theory, we would need to find evidence that L2 readers use of English affects their emotional responses to specific words.

Our data shows that this is precisely the case - words which are used by our L2 readers (language used in a university setting) showed less difference between L1 and L2 ratings than words that are seldom used by our L2 readers (words which they would use at home with their family). This supports the hypothesis that a word's emotionality is not only encoded by knowing the meaning of the word, but interacting with that word in a second language.

Furthermore, we found that length of time spent in an English-speaking country significantly affected how L2 raters responded to words. We found that the longer our L2 readers spent in Canada, and thus, the more sensorimotor interaction they have had with English, the more similar they were to native speakers of English. More specifically, an increase in real-world L2 exposure appears to gradually offset the attenuation of emotional responses characteristic of non-native use of L2, and engenders a more intensive, native-like emotional engagement with L2 lexical material. Thus, we conclude that word meaning and emotionality is created and stored as complex mental representations built by a wide variety of codings, including verbal and sensorimotor encodings.

These findings are in agreement with other studies which have found that the more everyday interactions a bilingual has with a given word or concept, the more emotional that word becomes. Puntoni, de Langhe, & van Osselaer (2009) found the more a concept is used in an L2 environment (rather than L1), the more emotional is becomes in a bilingual's second language.

These findings shed light on classic questions in bilingual research which have been previously difficult to disentangle. Although we were unable to determine whether proficiency or embodiment impacted L2 ratings, we hope that researchers can make use of the norms we are presenting, both in experimental work, but also to be used for further data mining.

Appendix

Table 1: Descriptive statistics of participants' English proficiency, including self-reported months spent in Canada, age English began in school, subjective English skills (1-low, 7-high), and percentage of time spent using English

	Mean	SD	Min	Max
Months in Canada	109.54	85.18	1.00	747.00
Age English taught in school	10.33	5.36	3.00	23.00
Subjective English listening skills	6.26	0.95	3.00	7.00
Subjective English speaking skills	6.04	1.16	2.00	7.00
Subjective English writing skills	5.94	1.24	3.00	7.00
Subjective English reading skills	6.17	1.02	3.00	7.00
Percentage of time using English	45.53	19.86	5.00	100.00

Table 2: Fixed effects of a regression model fitted to the L1-L2 difference in valence ratings, with a critical interaction of valence by subjective English reading proficiency (EngReadN) and a main effect of log transformed word frequency (lFreq). cValence stands for centered L1 valence. By-item SD = 0.433, by-participant SD = 0.242, residual SD = 2.66.

	Estimate	Std. Error	df	t value	Pr(> |t|)
(Intercept)	0.61	0.02	278	2.71	0.007
cValence	0.54	0.003	40740	17.80	<0.001
EngReadN	-0.01	0.003	210.6	-0.36	0.72
lFreq	-0.05	0.01	1986	-3.98	<0.001
cValence:EngReadN	-0.06	<0.001	67540	-13.841	<0.001

Table 3: Fixed effects of a regression model fitted to the L1-L2 difference in valence ratings, with a critical interaction of valence by square-root transformed time lived in Canada (sqrtmonth). cValence stands for centered L1 valence. By-item SD = 0.663, by-participant SD = 0.489, residual SD = 2.662.

	Estimate	Std. Error	df	t value	Pr(> |t|)
(Intercept)	0.15	0.1	222.9	0.70	0.48
cValence	0.27	0.02	7465	15.72	<0.001
sqrtmonth	0.017	0.01	210.9	1.68	0.09
cValence:sqrtmonth	-0.012	0.001	67780	-9.99	<0.001

References

Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review*, *96*(2), 358–372. https://doi.org/10.1037/0033-295X.96.2.358

Albaum, G., Best, R., & Hawkins, D. E. L. (1981). Continuous vs Discrete Semantic Differential Rating Scales. *Psychological Reports*, *49*, 83–86.

Allen, N. B., Trinder, J., & Brennan, C. (1999). Affective Startle Modulation in Clinical Depression: Preliminary Findings. *Biological Psychiatry*, *46*(4), 542–550. https://doi.org/https://doi.org/10.1016/S0006-3223(99)00025-6

Altarriba, J., & Bauer, L. M. (2004). The Distinctiveness of Emotion Concepts: A Comparison between Emotion, Abstract, and Concrete Words. *The American Journal of Psychology*, *117*(3), 389–410.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (DSM-5®). *American Psychiatric Pub*.

Anooshian, L. J., & Hertel, P. T. (1994). Emotionality in free recall: Language specificity in bilingual memory. *Cognition and Emotion*, *8*(6), 503–514. https://doi.org/10.1080/02699939408408956

Ayçiçeği-Dinn, A., & Caldwell-Harris, C. L. (2009). Emotion-memory effects in bilingual speakers: A levels-of-processing approach*. *Bilingualism: Language and Cognition*, *12*(03), 291. https://doi.org/10.1017/S1366728909990125

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(04), 577–660. https://doi.org/10.1017/S0140525X99002149

Batson, C. D., Early, S., & Salvarani, G. (1997). Perspective Taking: Imagining How Another Feels Versus Imaging How You Would Feel. *Personality and Social Psychology Bulletin*. https://doi.org/10.1177/0146167297237008

Bayer, M., Sommer, W., & Schacht, A. (2011). Emotional words impact the mind but not the body: Evidence from pupillary responses. *Psychophysiology*, *48*(11), 1554–1562. https://doi.org/10.1111/j.1469-8986.2011.01219.x

Betella, A., & Verschure, P. F. M. J. (2016). The affective slider: A digital self-assessment scale for the measurement of human emotions. *PLoS ONE*, *11*(2), 1–11. https://doi.org/10.1371/journal.pone.0148037

Bonin, P., Aubert, L., Malardier, N., & Niedenthal, P. M. (2016). Subjective Et De Valence Émotionnelle Pour 866 Mots, *103*, 655–694.

Bradley, M. M., & Lang, P. P. J. (1999). Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings. *Psychology*, *Technical*(C–1), 0. https://doi.org/10.1109/MIC.2008.114

Brase, J., & Mani, N. (2017). Effects of learning context on the acquisition and processing of emotional words in bilinguals. *Emotion*, *17*(4), 628–639. https://doi.org/10.1037/emo0000263

Briesemeister, B. B., Kuchinke, L., & Jacobs, A. M. (2011). Discrete emotion effects on lexical decision response times. *PLoS ONE*, *6*(8). https://doi.org/10.1371/journal.pone.0023743

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990. https://doi.org/10.3758/BRM.41.4.977

Bylsma, L. M., Morris, B. H., & Rottenberg, J. (2008). A meta-analysis of emotional reactivity in major depressive disorder. *Clinical Psychology Review*, *28*(4), 676–691. https://doi.org/10.1016/j.cpr.2007.10.001

Caldwell-Harris, C. L. (2015). Emotionality Differences Between a Native and Foreign Language. *Current Directions in Psychological Science*, *24*(3), 214–219. https://doi.org/10.1177/0963721414566268

Caldwell-Harris, C. L., Tong, J., Lung, W., & Poo, S. (2011). Physiological reactivity to emotional phrases in Mandarin--English bilinguals. *International Journal of Bilingualism*, *15*(3), 329–352. https://doi.org/10.1177/1367006910379262

Canli, T., Sivers, H., Whitfield, S. L., Gotlib, I. H., & Gabrieli, J. D. E. (2002). Amygdala response to happy faces as a function of extraversion. *Science*, *296*(5576), 2191. https://doi.org/10.1126/science.1068749

Cartwright, C., Gibson, K., Read, J., Cowan, O., & Dehar, T. (2016). Long-term antidepressant use: patient perspectives of benefits and adverse effects. *Patient Preference and Adherence*, *Volume 10*, 1401–1407. https://doi.org/10.2147/PPA.S110632

Carver, C. S., & White, T. L. (1994). Behavioral Inhibition, Behavioral Activation and affective responses to impending reward and punishment. *Journal of Personality and Social Psychology*, *67*, 319–333.

Chen, M., & Bargh, J. A. (1997). Nonconscious behavioral confirmation processes: The self-fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology*, *33*(5), 541–560. https://doi.org/10.1006/jesp.1997.1329

Citron, F. M. M., Weekes, B. S., & Ferstl, E. C. (2013). Effects of valence and arousal on written word recognition: Time course and ERP correlates. *Neuroscience Letters*, *533*, 90–95.

Clark, J., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*(3), 149–210. https://doi.org/10.1007/BF01320076

Cutrona, C. E., & Troutman, B. R. (1986). Social Support , Infant Temperament , and Parenting Self-Efficacy : A Mediational Model of Postpartum Depression Author (s): Carolyn E . Cutrona and Beth R . Troutman Published by : Wiley on behalf of the Society for Research in Child Development Stable, *57*(6), 1507–1518.

Daneman, M., & Carpenter, P. A. (1980). Individual Differences in Working Memory and Reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466.

De Grauwe, S., Willems, R. M., Rueschemeyer, S. A., Lemhöfer, K., & Schriefers, H. (2014). Embodied language in first- and second-language speakers: Neural correlates of processing motor verbs. *Neuropsychologia*, *56*(1), 334–349. https://doi.org/10.1016/j.neuropsychologia.2014.02.003

De Pascalis, V., Fiore, A. D., & Sparita, A. (1996). Personality, event-related potential (ERP) and heart rate (HR): An investigation of Gray’s theory. *Personality and Individual Differences*, *20*(6), 733–746. https://doi.org/10.1016/0191-8869(96)00016-5

De Pascalis, V., Strippoli, E., Riccardi, P., & Vergari, F. (2004). Personality, event-related potential (ERP) and heart rate (HR) in emotional word processing. *Personality and Individual Differences*, *36*(4), 873–891. https://doi.org/10.1016/S0191-8869(03)00159-4

Decety, J., & Batson, C. D. (2007). Social neuroscience approaches to interpersonal sensitivity. *Social Neuroscience*, *2*(3–4), 151–157. https://doi.org/10.1080/17470910701506060

Degner, J., Doycheva, C., & Wentura, D. (2012). It matters how much you talk: On the automaticity of affective connotations of first and second language words. *Bilingualism*, *15*(1), 181–189. https://doi.org/10.1017/S1366728911000095

Dewaele, J.-M. (2004). The Emotional Force of Swearwords and Taboo Words in the Speech of Multilinguals. *Journal of Multilingual and Multicultural Development*, *25*(23), 204–222. https://doi.org/10.1080/01434630408666529

Dewaele, J. M., & Nakano, S. (2013). Multilinguals' perceptions of feeling different when switching languages. *Journal of Multilingual and Multicultural Development*, *34*(2), 107–120. https://doi.org/10.1080/01434632.2012.712133

Dresler, T., Mériau, K., Heekeren, H. R., & Van Der Meer, E. (2009). Emotional Stroop task: Effect of word arousal and subject anxiety on emotional interference. *Psychological Research*, *73*(3), 364–371. https://doi.org/10.1007/s00426-008-0154-6

Ehrman, M. E., Leaver, B. Lou, & Oxford, R. L. (2003). A brief overview of individual differences in second language learning. *System*, *31*(3), 313–330. https://doi.org/10.1016/S0346-251X(03)00045-9

Eilola, T. M., Havelka, J., & Sharma, D. (2007). Emotional activation in the first and second language. *Cognition and Emotion*, *21*(5), 1064–1076. https://doi.org/10.1080/02699930601054109

Eisman, A. B., Stoddard, S. a, Heinze, J., Caldwell, C. H., & Zimmerman, M. a. (2015). Depressive symptoms, social support, and violence exposure among urban youth: A longitudinal study of resilience. *Developmental Psychology*, *51*(9), 1307–1316. https://doi.org/10.1037/a0039501

Ekman, P. (1992). An Argument for Basic Emotions. *Cognition and Emotion*, *6*(3–4), 169–200. https://doi.org/10.1080/02699939208411068

Elliot, A. J. (2006). The Hierarchical Model of Approach-Avoidance Motivation. *Motivation and Emotion*, *30*(2), 111–116. https://doi.org/10.1007/s11031-006-9028-7

Epstein, J., Pan, H., Kocsis, J., Yang, Y., Butler, T., Chusid, J., … Silbersweig, D. (2006). Lack of Ventral Striatal Response to Positive Stimuli inDepressed Versus Normal Subjects. *The American Journal of Psychiatry*, 1784. https://doi.org/10.1176/ajp.2006.163.10.1784

Farrow, T. F. D., & Woodruff, P. W. R. (2008). Empathy in Mental Illness. *Psychiatric Services*, *59*, 581–581. https://doi.org/10.1176/appi.ps.59.5.581

Feldman Barrett, L., & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology*, *74*(4), 967–984. https://doi.org/10.1037/0022-

3514.74.4.967

Ferré, P., García, T., Fraga, I., Sánchez-Casas, R., & Molero, M. (2010). Memory for emotional words in bilinguals: Do words have the same emotional intensity in the first and in the second language? *Cognition and Emotion*, *24*(5), 760–785. https://doi.org/10.1080/02699930902985779

Foroni, F. (2015). Do we embody second language? Evidence for 'partial' simulation during processing of a second language. *Brain and Cognition*, *99*, 8–16. https://doi.org/10.1016/j.bandc.2015.06.006

George, L., Blazer, D., Hughes, D., & Fowler, N. (1989). Social support and the outcome of major depression. *British Journal of Psychiatry*, *154*, 478–485.

Golin, S., Hartman, S. A., Klatt, E. N., Munz, K., & Wolfgang, G. L. (1977). Effects of Self-esteem Manipulation on Arousal and Reactions to Sad Models in Depressed and Nondepressed College Students. *Journal of Abnormal Psychology*, *86*(4), 435–439.

Haaga, D. A. F., & Beck, A. T. (1995). Perspectives on depressive realism: Implications for cognitive theory of depression. *Behaviour Research and Therapy*, *33*(1), 41–48. https://doi.org/10.1016/0005-7967(94)E0016-C

Harris, C. L., Ayçiçeği, A., & Gleason, J. B. (2003). Taboo words and reprimands elicit greater autonomic reactivity in a first language than in a second language. *Applied Psycholinguistics*, *24*(4), 561–579. https://doi.org/10.1017/S0142716403000286

Harris, C. L., Gleason, J. B., & Ayçiçeği, A. (2006). When is a First Language More Emotional? Psychophysiological Evidence from Bilingual Speakers. In *Bilingual minds: Emotional experience, expression and representation* (pp. 257–283). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.69.5830%0Ahttp://www.zeit.de/wirtschaft/2014-08/russland-import-sanktionen%0Ahttp://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.69.5830%0Ahttp://www.zeit.de/wirtschaft/2014-08/russland-import-sa

Hellman, A. B. (2011). Vocabulary size and depth of word knowledge in adult-onset second language acquisition. *International Journal of Applied Linguistics*, *21*(2), 162–182. https://doi.org/10.1111/j.1473-4192.2010.00265.x

Hinojosa, J. a., Martínez-García, N., Villalba-García, C., Fernández-Folgueiras, U., Sánchez-Carmona, a., Pozo, M. a., & Montoro, P. R. (2015). Affective norms of 875 Spanish words for five discrete emotional categories and two emotional dimensions. *Behavior Research Methods*, (August).

https://doi.org/10.3758/s13428-015-0572-5

Imbault, C., Shore, D., & Kuperman, V. (2017). Reliability of the sliding scale for collecting affective responses to words. *Behavior Research Methods*.

Inoue, Y., Yamada, K., & Kanba, S. (2006). Deficit in theory of mind is a risk for relapse of major depression. *Journal of Affective Disorders*, *95*(1–3), 125–127. https://doi.org/10.1016/j.jad.2006.04.018

Inoue, Yumiko, Tonooka, Y., Yamada, K., & Kanba, S. (2004). Deficiency of theory of mind in patients with remitted mood disorder. *Journal of Affective Disorders*, *82*(3), 403–409. https://doi.org/10.1016/j.jad.2004.04.004

Itkes, O., Eviatar, Z., & Kron, A. (2019). Semantic and affective manifestations of ambi (valence). *Cognition and Emotion*, *9931*. https://doi.org/10.1080/02699931.2018.1564249

Itkes, O., Kimchi, R., Kron, A., & Carmel, M. (2017). Dissociating Affective and Semantic Valence, *146*(7), 1–36. https://doi.org/10.1037/xge0000291

Itzhak, I., Vingron, N., Baum, S. R., & Titone, D. (2017). Bilingualism in the real world: How proficiency, emotion, and personality in a second language impact communication in clinical and legal settings. *Translational Issues in Psychological Science*, *3*(1), 48–65. https://doi.org/10.1037/tps0000103

Johnson-Laird, P. N., & Oatley, K. (1989). The Language of Emotions: An Analysis of a Semantic Field. *Cognition and Emotion*, *3*(2), 81–123. https://doi.org/10.1080/02699938908408075

Jolliffe, D., & Farrington, D. P. (2006). Development and validation of the Basic Empathy Scale. *Journal of Adolescence*, *29*(4), 589–611. https://doi.org/10.1016/j.adolescence.2005.08.010

Just, M. A., & Carpenter, P. A. (1999). A Capacity Theory of Comprehension : Indicidual Dirrefences in Working Memory, (January), 1–16.

Kazanas, S. A., & Altarriba, J. (2016). Emotion Word Processing: Effects of Word Type and Valence in Spanish–English Bilinguals. *Journal of Psycholinguistic Research*, *45*(2), 395–406. https://doi.org/10.1007/s10936-015-9357-3

Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual Differences in Language Acquisition and Processing. *Trends in Cognitive Sciences*, *22*(2), 154–169. https://doi.org/10.1016/j.tics.2017.11.006

Kleinginna, P. R., & Kleinginna, A. M. (1981). A categorized list of motivation definitions, with a suggestion for a consensual definition. *Motivation and Emotion*, *5*(3), 263–291. https://doi.org/10.1007/BF00993889

Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The Representation of Abstract Words: Why Emotion Matters. *Journal of Experimental Psychology: General*, *140*(1), 14–34. https://doi.org/10.1037/a0021446

Kousta, S. T., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition*, *112*(3), 473–481. https://doi.org/10.1016/j.cognition.2009.06.007

Kuchinke, L., Võ, M. L. H., Hofmann, M., & Jacobs, A. M. (2007). Pupillary responses during lexical decisions vary with word frequency but not emotional valence. *International Journal of Psychophysiology*, *65*(2), 132–140. https://doi.org/10.1016/j.ijpsycho.2007.04.004

Kunzmann, U., & Grühn, D. (2005). Age differences in emotional reactivity: The sample case of sadness. *Psychology and Aging*, *20*(1), 47–59. https://doi.org/10.1037/0882-7974.20.1.47

Kuperman, V., Estes, Z., Brysbaert, M., & Warriner, A. B. (2014). Emotion and language: Valence and arousal affect word recognition. *Journal of Experimental …*, (4). Retrieved from http://psycnet.apa.org/psycarticles/2014-03909-001

Kuperman, V., & Van Dyke, J. a. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language*, *65*(1), 42–73. https://doi.org/10.1016/j.jml.2011.03.002

Kuppens, P., Tuerlinckx, F., Yik, M., Koval, P., Coosemans, J., Zeng, K. J., & Russell, J. A. (2017). The Relation Between Valence and Arousal in Subjective Experience Varies With Personality and Culture. *Journal of Personality*, *85*(4), 530–542. https://doi.org/10.1111/jopy.12258

Kuznetsova, A., Brockhoff, P., & Christensen, R. (2016). lmerTest: Tests in Linear Mixed Effects Models. *R Package Version*, *3.0.0*, https://cran.r-project.org/package=lmerTest. Retrieved from https://cran.r-project.org/package=lmerTest

Lang, P. J. (1995). The Emotion Probe: Studies of Motivation and Attention. *American Psychologist*. https://doi.org/10.1037/0003-066X.50.5.372

Larsen, R. J., & Ketelaar, T. (1991). Personality and susceptibility to positive and negative emotional states. *Journal of Personality and Social Psychology*, *61*(1), 132–140. https://doi.org/10.1037/0022-

3514.61.1.132

Levenson, R. W., Carstensen, L. L., Friesen, W. V., & Ekman, P. (1991). Emotion, physiology, and expression in old age. *Psychology and Aging*, *6*(1), 28–35. https://doi.org/10.1037/0882-7974.6.1.28

Lewinsohn, P. M., Lobitz, W. C., & Wilson, S. (1973). “Sensitivity” of Depressed Individuals to Aversive Stimuli. *Journal of Abnormal Psychology, 81*(3), 259–263.

Long, L. S., & Titone, D. A. (2007). Psychopathy and verbal emotion processing in non-incarcerated males. *Cognition and Emotion*, *21*(1), 119–145. https://doi.org/10.1080/02699930600551766

Mackay, D. G., Shafto, M., Taylor, J. K., Marian, D. E., Abrams, L., & Dyer, J. R. (2004). Relations between emotion, memory, and attention: Evidence from taboo Stroop, lexical decision, and immediate memory tasks. *Memory & Cognition*, *32*(3), 474–488. https://doi.org/10.3758/BF03195840

Macleod, J. W., Lawrence, M. a, McConnell, M. M., Eskes, G. a, Klein, R. M., & Shore, D. I. (2010). Appraising the ANT: Psychometric and theoretical considerations of the Attention Network Test. *Neuropsychology*, *24*(5), 637–651. https://doi.org/10.1037/a0019803

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research : JSLHR*, *50*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

Mcwhirter, B. T., Besett-Alesch, T. M., Horibata, J., & Gat, I. (2002). Loneliness in High Risk Adolescents: The Role of Coping, Self-Esteem, and Empathy. *Journal of Youth Studies*, *5*(1), 69–84. https://doi.org/10.1080/1367626012011177

Méndez-Bértolo, C., Pozo, M. A., & Hinojosa, J. A. (2011). Word frequency modulates the processing of emotional words: Convergent behavioral and electrophysiological data. *Neuroscience Letters*, *494*(3), 250–254. https://doi.org/10.1016/j.neulet.2011.03.026

Monnier, C., & Syssau, A. (2014). Affective Norms for French Words (FAN). *Behavior Research Methods*, *46*(4), 1128–1137. https://doi.org/10.3758/s13428-013-0431-1

Moors, A., De Houwer, J., Hermans, D., Wanmaker, S., van Schie, K., Van Harmelen, A.-L., … Brysbaert, M. (2013). Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words. *Behavior Research Methods*, *45*(1), 169–177. https://doi.org/10.3758/s13428-012-0243-8

Mor, N., & Winquist, J. (2002). Self-focused attention and negative affect: A meta-analysis. *Psychological Bulletin*, *128*(4), 638–662. https://doi.org/10.1037/0033-2909.128.4.638

O'Connor, L. E., Berry, J. W., Lewis, T., Mulherin, K., & Crisotomo, P. S. (2007). Empathy and depression: the moral system on overdrive. *Empathy in Mental Ilness*. https://doi.org/10.1017/CBO9780511543753.005

Orne, M. T. (1962). On the Social Psychology of Psychological Experiment: With Particular Reference to Demand Characteristics and Their Implications. *American Psychologist*, *17*(11), 776–783. Retrieved from https://journals.scholarsportal.info/details/0003066x/v17i0011/776_otspottdcati.xml

Ostaszewski, K., & Zimmerman, M. A. (2006). The effects of cumulative risks and promotive factors on urban adolescent alcohol and other drug use: A longitudinal study of resiliency. *American Journal of Community Psychology*, *38*(3–4), 237–249. https://doi.org/10.1007/s10464-006-9076-x

Pearson, C., Janz, T., & Ali, J. (2013). *Mental and substance use disorders in Canada*.

Peirce, R., Frone, M., Russell, M., Cooper, M., & Mudar, P. (2000). A Longitudinal Model of Social Contact, Social Support, Depression, and Alcohol Use. *Health Psychol*, *19*(1), 28–38. https://doi.org/10.1037/0278-6133.19.1.28

Pennebaker, J., Booth, R., & Francis, M. (2007). Linguistic inquiry and word count: LIWC. Austin, TX: liwc net.

Perfetti, C. (2007). Reading Ability: Lexical Quality to Comprehension. *Scientific Studies of Reading*, *11*(4), 357–383. https://doi.org/10.1080/10888430701530730

Ponari, M., Rodriguez-Cuadrado, S., & Vinson, D. (2014). Processing advantage for emotional words in bilingual speakers. *Emotion*, *15*(January), 644–652. https://doi.org/10.1037/emo0000061

Ponari, M., Rodriguez-Cuadrado, S., Vinson, D., Fox, N., Costa, A., & Vigliocco, G. (2014). Processing advantage for emotional words in bilingual speakers. *Emotion*, *15*(January), 644–652. https://doi.org/10.1037/emo0000061

Pratto, F., & John, O. P. (1991). Automatic Vigilance: The Attention-Grabbing Power of Negative Social Information. *Journal of Personality and Social Psychology*, *61*(3), 380–391. https://doi.org/10.1037/0022-3514.61.3.380

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *The Behabioral and Brain Sciences*, *4*, 515–526. https://doi.org/10.1016/j.celrep.2011.1011.1001.7.

Puntoni, S., de Langhe, B., & van Osselaer, S. M. J. (2009). Bilingualism and the Emotional Intensity of Advertising Language. *Journal of Consumer Research*, *35*(6), 1012–1025. https://doi.org/10.1086/595022

Pyszczynski, T., Holt, K., & Greenberg, J. (1987). Depression, self-focused attention, and expectancies for positive and negative future life events for self and others. *Journal of Personality and Social Psychology*, *52*, 994–1001. https://doi.org/10.1037/0022-3514.52.5.994

R Core Team, I. (2017). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna (2014). R Foundation for Statistical Computing Authors R Core Team.

Recio, G., Conrad, M., Hansen, L. B., & Jacobs, A. M. (2014). On pleasure and thrill: The interplay between arousal and valence during visual word recognition. *Brain and Language*, *134,* 34–43. https://doi.org/10.1016/j.bandl.2014.03.009

Redondo, J., Fraga, I., Padrón, I., & Comesaña, M. (2007). The Spanish adaptation of ANEW (affective norms for English words). *Behavior Research Methods*, *39*(3), 600–605. https://doi.org/10.3758/BF03193031

Rickard, N. S. (2004). Intense emotional responses to music: a test of the physiological arousal hypothesis. *Psychology of Music*, *32*(4), 371–388. https://doi.org/10.1177/0305735604046096

Rottenberg, J., & Gotlib, I. (2004). Socioemotional Functioning in Depression. In *Mood Disorders: A Handbook of Science and Practice* (pp. 61–77).

Rottenberg, J., Gross, J. J., & Gotlib, I. H. (2005). Emotion context insensitivity in major depressive disorder. *Journal of Abnormal Psychology*, *114*(4), 627–639. https://doi.org/10.1037/0021-843X.114.4.627

Rottenberg, J., & Hindash, A. C. (2015). Emerging evidence for emotion context insensitivity in depression. *Current Opinion in Psychology*, *4*, 1–5. https://doi.org/10.1016/j.copsyc.2014.12.025

Rush, J., Trivedi, M. H., Ibrahim, H. M., Carmody, T. J., Arnow, B., Klein, D. N., … Keller, M. B. (2003). The 16-Item Quick Inventory of Depressive Symptomatology. *Depression*, *54*(5), 573–583. https://doi.org/10.1016/S0006-3223(03)01866-8

Russell, C. J., & Bobko, P. (1992). Moderated Regression Analysis and Likert Scales Too Coarse for Comfort. *Journal of Applied Psychology*, *77*(3), 336–342.

Russell, J. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*. Retrieved from http://psycnet.apa.org/psycinfo/1981-25062-001

Sandell, K., & Bornäs, H. (2017). Functioning Numbness Instead of Feelings as a Direction: Young Adults' Experiences of Antidepressant Use. *Sociology*, *51*(3), 543–558. https://doi.org/10.1177/0038038515591947

Santiago-Rivera, A. L., & Altarriba, J. (2002). The Role of Language in Therapy With the Spanish-English Bilingual Client. *Professional Psychology: Research and Practice*, *33*(1), 30–38. https://doi.org/10.1037/0735-7028.33.1.30

Schutte, N. S., Malouff, J. M., Bobik, C., Coston, T. D., Greeson, C., Jedlicka, C., … Wendorf, G. (2001). Emotional intelligence and interpersonal relations. *Journal of Social Psychology*, *141*(4), 523–536. https://doi.org/10.1080/00224540109600569

Scott, G., O'Donnell, P., Leuthold, H., & Sereno, S. (2009). Early emotion word processing: evidence from event-related potentials. *Biological Psychology*, *80*(1), 95–104. https://doi.org/S0301-0511(08)00073-2 [pii]\r10.1016/j.biopsycho.2008.03.010

Scott, G., O'Donnell, P., & Sereno, S. (2012). Emotion words affect eye fixations during reading. *Journal of Experimental Psychology: Learning Memory and Cognition*, *38*(3), 783–792. https://doi.org/10.1037/a0027209

Segalowitz, N., Trofimovich, P., Gatbonton, E., & Sokolovskaya, A. (2008). Feeling affect in a second language: The role of word recognition automaticity. *The Mental Lexicon*, *3*(1), 47–71. https://doi.org/10.1075/ml.3.1.05seg

Sharot, T. (2011). The optimism bias. *Current Biology*, *21*(23), R941–R945. https://doi.org/10.1016/j.cub.2011.10.030

Sheikh, N. A., & Titone, D. A. (2013). Sensorimotor and linguistic information attenuate emotional word processing benefits: An eye-movement study. *Emotion*, *13*(6), 1107–1121. https://doi.org/10.1037/a0032417

Sheikh, N., & Titone, D. (2015). The embodiment of emotional words in a second language: An eye-movement study. *Cognition & Emotion*, (April), 1–13. https://doi.org/10.1080/02699931.2015.1018144

Shields, M., Gorber, S. C., Janssen, I., & Tremblay, M. S. (2011). Bias in self-reported estimates of obesity in Canadian health surveys: An update on correction equations for adults. *Health Reports*, *22*(3).

Sigmon, S. T., & Nelson-Gray, R. O. (1992). Sensitivity to Aversive Events in Depression: Antecedent, Concomitant, or Consequent? *Journal of Psychopathology and Behavioral Assessment*, *14*(3), 225–246.

Skehan, P. (1991). Individual Differences in Second Language Learning. *Studies in Second Language Acquisition*, *13*(02), 275. https://doi.org/10.1017/S0272263100009979

Sloan, D. M., Strauss, M. E., Quirk, S. W., & Sajatovic, M. (1997). Subjective and expressive emotional responses in depression. *Journal of Affective Disorders*, *46*, 135–141. https://doi.org/https://doi.org/10.1016/S0165-0327(97)00097-9

Sloan, D. M., Strauss, M. E., & Wisner, K. L. (2001). Diminished Response to Pleasant Stimuli by Depressed Women. *Journal of Abnormal Psychology*, *110*(3), 488–493. https://doi.org/I0.1037//0021-843X.110.3.488

Smith, D. P., Hillman, C. H., & Duley, A. R. (2005). Influences of Age on Emotional Reactivity During Picture Processing. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *60*(1), P49–P56. https://doi.org/10.1093/geronb/60.1.P49

Soares, A. P., Comesaña, M., Pinheiro, A. P., Simões, A., & Frade, S. (2012). Behavior and Research Methods Corresponding author : *Behavior Research Methods*, *44*(1), 256–269. https://doi.org/10.3758/s13428-011-0131-7

Stadthagen-González, H., Ferré, P., Pérez-Sánchez, M. A., Imbault, C., & Hinojosa, J. A. (2017). Norms for 10,491 Spanish words for five discrete emotions: Happiness, disgust, anger, fear, and sadness. *Behavior Research Methods*. https://doi.org/10.3758/s13428-017-0962-y

Stadthagen-Gonzalez, H., Imbault, C., Pérez Sánchez, M. A., & Brysbaert, M. (2017). Norms of valence and arousal for 14,031 Spanish words. *Behavior Research Methods*, 1–45. https://doi.org/10.3758/s13428-015-0700-2

Statistics Canada. (2013). *National Household Survey*. *Immigration and Ethnocultural Diversity in Canada* (Vol. 99-010-X20). https://doi.org/99-010-X2011001

Sutton, T. M., Altarriba, J., Gianico, J. L., & Basnight-Brown, D. M. (2007). The automatic access of emotion: Emotional Stroop effects in Spanish-English bilingual speakers. *Cognition and Emotion*, *21*(5), 1077–1090. https://doi.org/10.1080/02699930601054133

Thoma, P., Norra, C., Juckel, G., Suchan, B., & Bellebaum, C. (2015). Performance monitoring and empathy during active and observational learning in patients with major depression. *Biological Psychology*, *109*, 222–231. https://doi.org/10.1016/j.biopsycho.2015.06.002

Tsai, J. L., Knutson, B., & Fung, H. H. (2006). Cultural variation in affect valuation. *Journal of Personality and Social Psychology*, *90*(2), 288–307. https://doi.org/10.1037/0022-3514.90.2.288

Velez-Uribe, I., & Rosselli, M. (2019). The auditory and visual appraisal of emotion-related words in Spanish–English bilinguals. *Bilingualism: Language and Cognition*, *22*(1), 30–46. https://doi.org/10.1017/S1366728917000517

Vinson, D., Ponari, M., & Vigliocco, G. (2014). How does emotional content affect lexical processing? *Cognition & Emotion*, *28*(4), 737–746. https://doi.org/10.1080/02699931.2013.851068

Võ, M. L. H., Jacobs, A. M., Kuchinke, L., Hofmann, M., Conrad, M., Schacht, A., & Hutzler, F. (2008). The coupling of emotion and cognition in the eye: Introducing the pupil old/new effect. *Psychophysiology*, *45*(1), 130–140. https://doi.org/10.1111/j.1469-8986.2007.00606.x

Wang, Y. guang, Wang, Y. qiang, Chen, S. lin, Zhu, C. yan, & Wang, K. (2008). Theory of mind disability in major depression with or without psychotic symptoms: A componential view. *Psychiatry Research*, *161*(2), 153–161. https://doi.org/10.1016/j.psychres.2007.07.018

Waris, O., Soveri, A., Ahti, M., Hoffing, R. C., Ventus, D., Jaeggi, S. M., … Laine, M. (2017). A latent factor analysis of working memory measures using large-scale data. *Frontiers in Psychology*, *8*(JUN), 1–14. https://doi.org/10.3389/fpsyg.2017.01062

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191–1207. https://doi.org/10.3758/s13428-012-0314-x

Warriner, A. B., Shore, D. I., Schmidt, L. A., Imbault, C. L., & Kuperman, V. (2017). Sliding into happiness: A new tool for measuring affective responses to words. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *71*(1), 71–88. https://doi.org/10.1037/cep0000112

Watson, D., Clark, L. A., & Carey, G. (1988). Positive and negative affectivity and their relation to anxiety and depressive disorders. *Journal of Abnormal Psychology*, *97*(3), 346–353. https://doi.org/10.1037/0021-843X.97.3.346

Weisleder, A., & Fernald, A. (2013). Talking to Children Matters: Early Language Experience

Strengthens Processing and Builds Vocabulary. *Psychological Science*, *24*(11), 2143–2152. https://doi.org/10.1177/0956797613488145

Westbury, C., Keith, J., Briesemeister, B. B., Hofmann, M. J., & Jacobs, A. M. (2015). Avoid violence, rioting, and outrage; approach celebration, delight, and strength: Using large text corpora to compute valence, arousal, and the basic emotions. *The Quarterly Journal of Experimental Psychology*, *0*(0), 1–24. https://doi.org/10.1080/17470218.2014.970204

Wickham, H., & Chang, W. (2016). Create elegant data visualisations using the grammar of graphics. https://doi.org/10.1093/bioinformatics/btr406

Williams, J. M., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin*, *120*(1), 3–24. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8711015

Winskel, H. (2013). The emotional Stroop task and emotionality rating of negative and neutral words in late Thai-English bilinguals. *International Journal of Psychology*. Taylor & Francis. https://doi.org/10.1080/00207594.2013.793800

Wolkenstein, L., Schönenberg, M., Schirm, E., & Hautzinger, M. (2011). I can see what you feel, but i can't deal with it: Impaired theory of mind in depression. *Journal of Affective Disorders*, *132*(1–2), 104–111. https://doi.org/10.1016/j.jad.2011.02.010

Wolter, B. (2001). COMPARING THE L1 AND L2 MENTAL LEXICON A Depth of Individual Word Knowledge Model. *Studies in Second Language Acquisition*, *23*(1), S0272263101001024. https://doi.org/10.1017/S0272263101001024

Zuroff, D., Koestner, R., Moskowitz, D. S., McBride, C., Marshall, M., & Bagby, M. (2007). Autonomous motivation for therapy: A new common factor in brief treatments for depression. *Psychotherapy Research*, *17*(2), 137–148. https://doi.org/10.1080/10503300600919380

Chapter 3

Abstract

Warriner, Shore, Schmidt, Imbault, and Kuperman (2017) have recently proposed a slider task in which participants move a manikin on a computer screen towards or further away from a word and the distance (in pixels) is a measure of the word's valence. Warriner et al. have shown this task to be more valid than the widely used rating task but have not examined the reliability of the new methodology. This paper investigates multiple aspects of the task reliability. In Experiment 1 (E1.1-1.6), we show that the sliding scale has high split-half reliability (r = 0.868 to 0.931). In Experiment 2, we also show that the slider task elicits consistent repeated responses within a single session (Experiment 2: r = 0.804), and across two sessions separated by one week (Experiment 3: r = 0.754). Overall, the slider task, in addition to having high validity, is highly reliable.

Keywords: valence; arousal; emotion; reliability

Introduction

Words evoke affective responses, which can be indexed through subjective ratings of valence. Norms for these responses are available in English (Bradley & Lang, 1999; Warriner et al., 2013), French (Bonin et al., 2016; Monnier & Syssau, 2014), Spanish (Hinojosa et al., 2015; Redondo, Fraga, Padrón, & Comesaña, 2007; Stadthagen-Gonzalez et al., 2017), Dutch (Moors et al., 2013) and other languages. All of these studies, dating back to 1999 (Bradley & Lang, 1999), have one method in common; they present words in isolation and instruct participants to evaluate their valence (from negative to positive) and arousal (from calm to excited) on a rating scale from 1 to 9 (or in some cases, 1 to 7). A number of criticisms of this methodology led us to develop a new methodology (Warriner et al., 2017), summarized below. The present paper evaluates the reliability of this sliding scale methodology.

Criticisms of the popular rating scale fall into two main categories (Warriner, et al., 2017; Westbury, Keith, Briesemeister, Hofmann, & Jacobs, 2014). The first class of criticism concerns the nature of the data collected. A rating scale with 9 (or 7) individual points produces an ordinal measure, whereas an interval measure is preferable for most statistical tests. In fact, most papers treat this ordinal scale as an interval scale for statistical analyses, which can lead to violations of assumptions. Related to this general criticism, the typical rating scale does not allow a fine-grained output—observers can only provide integer responses, which can fail to capture the subtle effects of emotion evoked by a word, or individual variability in affective behaviour.

The second class of criticism concerns the necessity, within a typical rating study, to anchor the affective evaluation. Depending on the words chosen for the anchors, observers can be biased to overemphasize some words and diminish the impact of other words (Westbury et al., 2015). Typically, the words chosen (e.g., “pleasant” and “unpleasant”) are too mild for words that raters encounter in these studies, such as “rapist”. Inappropriate anchoring can lead to drift across the study once extreme words are encountered. Ideally, the scale should allow observers to maximize the range of values used without reliance on the specific words presented.

To counter these concerns, we developed the *slider* task (Warriner et al., 2017). A humanoid manikin is placed in the centre of a vertical line, with a word at the top or the bottom of the line. Participants, who are instructed that the manikin represents themselves, can slide the manikin as close to or as far away from the word as they prefer. The distance from the word represents the participants' affective response to a word; a greater distance indicates lower valence (more negative) and a smaller distance indicates higher valence (more positive). Distance is a continuous interval-scale variable, which is in practice discretized into a number of pixels[1] on a computer screen, just like the continuous variable of time is discretized into fractions of seconds by chronometric instruments. Warriner et al. (2017) found that distance was negatively correlated with word valence in a variety of populations (undergraduates in lab: r = -0.62; adults online: r = -0.58). Additionally, Warriner et al. (2017) found that the slider task is sensitive to individual differences; those who are shyer tend to position the manikin farther from all stimuli than those who are less shy (56 pixels or 10% of the scale). As well, the slider is sensitive to gender differences; females position the manikin closer to words that are rated as being more positive by female than male raters, and the same is true for males. The range of 600 pixels that we utilize in the slider task also allows researchers to capture subtle individual differences that were lost in a smaller 9 point ordinal scale (e.g., the subtle tendency of relatively sociable individuals to keep a shorter distance to all words; (Warriner et al., 2017). These findings are in line with an earlier proposal to move from discrete ratings scale to a continuous slider by Albaum, Best, and Hawkins (1981). They stated that a discrete ratings scale and a continuous scale produced similar aggregate data, but a continuous scale allowed for greater discrimination at the individual level.

Finally, this task has no anchors and removes the mention of valence from task instructions. By tapping into implicit approach–avoidance tendencies, the task avoids explicit linkages to artificial valence terminology (e.g., "pleasant" and "unpleasant"). The lack of traditional semantic labels alluding to valence may alter the psychological construct measured by the slider task, and may cause participants to tap into approach-avoidance behaviour instead of producing affective responses (see the General Discussion below and the Future Directions in Warriner et al., 2017). An argument can also be made that the lack of anchors may cause participants to be confused and not perform the task as intended. Although possible, it is unlikely that participants are confused or engage in a different behavioral pattern. Warriner et al. (2017) administered the

[1] We thank our reviewers for bringing double density displays to our attention. In future studies, we will refer to "steps" as our dependent variable of interest, rather than "pixels" which may be an inaccurate measurement on double density displays.

slider task with and without anchors and found that there was no difference in the functional relationship between valence and distance. In sum, the slider task provides a new method of collecting affective ratings that is more valid than past methods. The utility of a similar Affective Slider method for measuring the valence and arousal of pictures taken from the International Affective Picture System (IAPS) has been demonstrated by Betella and Verschure (2016). Given the validity of the task (Warriner et al., 2017), we expect many researchers and clinicians to be interested in the assessments of emotion that come from this task; however, before we can evaluate individual and group differences in emotional responses, we must ask about its reliability. The present paper evaluates the reliability of the slider task through several converging methods: Experiment 1 utilizes previously collected data and applies a split-half analysis (cf. Macleod et al., 2010); Experiment 2 collects new data in a repeated-measures design within a single data collection session; and Experiment 3 uses a standard test–retest design from two sessions separated by a week. In all cases, the measure of interest was the slope of the best fitting regression line between distance from the word (cf. Warriner et al., 2017) as the dependent measure and the normed rated valence of the word (cf. Warriner et al., 2013) as the independent variable. To be specific, we assessed to what extent the slope of this line from one sample predicted the slope from a second sample within the same individual. With a split-half analysis, the two samples came from the same session (Experiment 1). With the repeated measures design, observers rated the same words in two separate blocks of trials in one session (Experiment 2); and, with the test-retest design the same words were rated on two separate occasions separated by one week (Experiment 3). We performed an additional analysis in Experiments 2 and 3 to remove the effect of valence, and instead measure the reliability of distance. We assessed to what extent the distance from the word in one sample predicted the distance from the word in a second sample within the same individual.

Experiment 1: Split Half Analysis

The split-half analysis utilized previously collected data (Warriner et al., 2017) to estimate the reliability of the slider task. We combined the standard split-half analysis with a resampling technique to enhance the robustness of our estimate and provide confidence intervals on our reliability estimate (cf. MacLeod et al., 2010). Each experiment presented the same 250 words drawn from separate quintiles of valence; there were 50 words rated for each of five different valence subranges. To determine split-half reliability, we randomly selected two samples of twenty words from each quintile without replacement, creating two groups of 100 words. For each group and each participant, we estimated a regression slope for distance as predicted by valence: two regression slopes were thus generated for each participant. Reliability was estimated by examining the correlation of these two slopes across participants. This sampling was conducted 10,000 times for each of the five experiments, four of which are reported in Warriner et al. (2017), and the fifth is data from the first session of the current paper's Experiment 3 (labeled here Experiments 1.1 – 1.5). The data from Experiment 2 was not included in the split-half analysis as Experiment 2 only presented half of the 250 words used in Experiments 1.1 – 1.5. These five experiments varied in experimental settings, task instructions, and participant demographics. Experiment 1.6 analyzed the combined data from all five experiments, weighing each experiment by the number of participants.

Method

Participants and Procedure. All experiments used a similar set of instructions and similar numbers of participants (See Table 1 for specifics from each experiment; see Warriner et al., 2017 for detailed methods). Each participant was seated in front of a computer monitor with a screen resolution of 1024x768 placed approximately 57cm in front of the participant (E1.4 and 1.5 used an online version of the task—the specifics of the monitor and computer varied across participants). Following a central fixation point, a humanoid manikin was presented at the centre of the computer monitor along a line with a single word at the top or bottom of the screen (see Figure 1). Participants moved the manikin up or down to position it at their preferred location as close to, or as far away from, the word. Experiments 1.1, 1.2, 1.3 and 1.4 all had the same instructions, which read:

> *[… On a] screen, you will see a word at the top of the screen with a vertical line below it. There will be a person in the centre of that line. The person represents you. Your job is to assess how close you would like to be to the word and communicate that by clicking a point on the line to position the person (you). For example, if the word was DISASTER, you'd probably want to be far away and would click somewhere on the line far away from the word. But if the word was TRIUMPH you might want to be close and would place the manikin somewhere on the line really close to the word. […]*

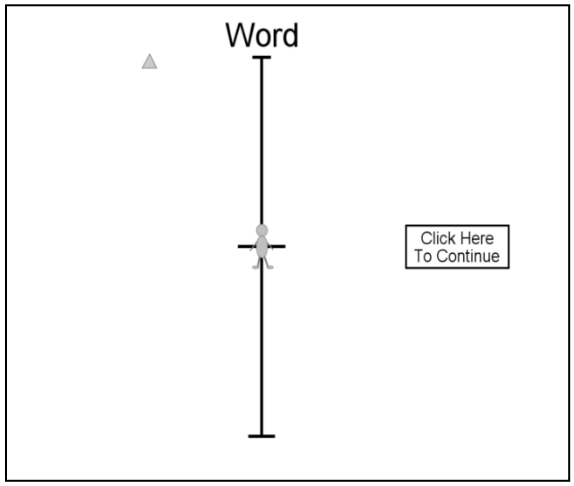

Figure 1. The slider scale: a humanoid manikin and word for each trial.

Experiment 1.5 had a slightly modified set of instructions where anchor words like *disaster* or *triumph* were not mentioned to the participant. Those instructions read:

> *[…] On each of the following screens, you will first see a plus sign in the centre. That's to center the mouse for the next screen. Click on the plus and you will see a word either at the top or the bottom with a vertical line below or above it. There will be a person in the centre of that line. The person represents you. You can move "yourself" closer to or further away from the word. Position yourself where you prefer to be. […]*

Stimuli The word set for each experiment was the same. It consisted of 250 monosyllabic words chosen from a set of 13,763 words that had previously been rated for valence and arousal (Warriner et al., 2013). The words were divided into 25 bins (5 quintiles of valence x 5 quintiles of arousal) with 10 words randomly drawn from each bin. Thus, there were 50 words at each of 5 levels of valence, which varied in their arousal levels. This ensured that valence and arousal were not correlated (p = -0.019), and thus in what follows we only measure any effects of valence and not arousal. The mean word length was 4.4 characters, and the mean natural-log SUBTLEX frequency was 6.3 (Brysbaert & New, 2009). We used natural-log frequency because frequency is exponentially represented, and the natural log makes this distribution closer to normal and easier to interpret.

Table 1. A summary of participants, procedure and critical results (correlation between valence and distance, and split-half reliability) of Experiments 1.1-1.5.

	Participants	Procedure	Split Half Results	Spearman Brown Correction
1.1	43 participants (35 female), ranging from 17 to 25 years of age (*M* = 19.07, *SD* = 1.98)	Participants at McMaster University took part in the study for partial course credit. No more than 10 participants at a time completed the study in a computer lab on campus. This data is taken from Warriner et al.'s (2017) Experiment 1.	ρ = 0.873, 95% CI [0.800, 0.926]	ρ = 0.932, 95% CI [0.889, 0.962]
1.2	30 participants (all female), ranging in age from 18 to 21 (*M* = 19, *SD* = 1.02)	Participants at McMaster University took part in the study for partial course credit. No more than 10 participants at a time completed the study in a computer lab on campus. Prior to the start of the experiment, participants were given 4 personality questionnaires to fill out. This data is taken from Warriner et al.'s (2017) Experiment 2.	ρ = 0.868, 95% CI [0.786, 0.929]	ρ = 0.929, 95% CI [0.880, 0.963]
1.3	36 participants (33 female) ranging in age from 18 to 27 (*M* = 18.69, *SD* = 1.59)	Participants at McMaster University took part in the study for partial course credit. No more than 10 participants at a time completed the study in a computer lab on campus. This data is taken from the first session of the present Experiment 3.	ρ = 0.883, 95% CI [0.812, 0.934]	ρ = 0.938, 95% CI [0.896, 0.966]

1.4	32 participants (13 female), ranging in age from 19 to 55 (*M* = 34.06, *SD* = 8.68)	An ad was posted on Amazon Mechanical Turk for participants to complete the study from their home for monetary compensation. All participants were based in the USA. This data is taken from Warriner et al.'s (2017) Experiment 3.	ρ = 0.931, 95% CI [0.886, 0.963]	ρ = 0.964, 95% CI [0.939, 0.981]
1.5	36 participants (13 female) ranging in age from 21 to 60 (*M* = 34.42, *SD* = 8.95)	An ad was posted on Amazon Mechanical Turk for participants to complete the study from their home for monetary compensation. All participants were based in the USA. Anchors were removed from the instructions given to participants. This data is taken from Warriner et al.'s (2017) Experiment 4.	ρ = 0.898, 95% CI [0.831, 0.945]	ρ = 0.946, 95% CI [0.908, 0.972]
1.6	177 participants (124 female), ranging in age from 17 to 60 (*M* = 24.81, *SD* = 9.31)	The data from experiments 1.1 through 1.5 were compiled together to create one data set.	ρ = 0.891, 95% CI [0.859, 0.917]	ρ = 0.942, 95% CI [0.924, 0.957]

Results and Discussion

The primary dependent variable for all experiments was the distance between the position of the manikin when the participant pressed the Submit button and the presented word, measured in pixels (range = 1 (closest to the word) – 600 (farthest away from the word)). The participants were able to move the manikin as many times as they wanted before clicking Submit—only the final location of the manikin was used in our data analysis. The independent variable of interest for all experiments presented here was the word's valence rating (Warriner et al., 2013), which varied between 1 (very unhappy) and 9 (very happy); each word was rated by at least 20 raters.

To assess the split half reliability of each experiment, we separated the 250 words in the stimulus list into 5 bins of valence, broken down by quantiles. We randomly selected 20 words (from approximately 50 words), without replacement, from each bin twice for each participant, to construct two separate groups of 100 words (20 words x 5 bins). In each group, we computed the slope (i.e., the beta coefficient) of the respective regression line between valence and distance. This resulted in 2 slope coefficients for each participant. We then found the correlation for these two slope estimates across all participants. We repeated this process 10,000 times which resulted in 10,000 correlations. The mean of these 10,000 correlations provided our estimate of the split half reliability; the distribution of correlations allowed us to estimate confidence intervals. We used R version 3.01 (R Core Development Team, 2013) to perform our statistical analyses in this and subsequent experiments.

The split half reliability for each of Experiments 1.1 – 1.5 ranged from 0.868 to 0.931 (see Table 1 and Figure 2), which can be considered a very high level of reliability. Experiment 1.6 combines the data from all 5 experiments and weighs each experiment by the number of participants. This analysis produced a reliability of 0.891 [95% CI = 0.859–0.917]. Thus, this new method of collecting affective ratings towards words provides a reliable measure of valence. This was found across multiple experiments with different participant pools, recruitment methods, and experimental settings.

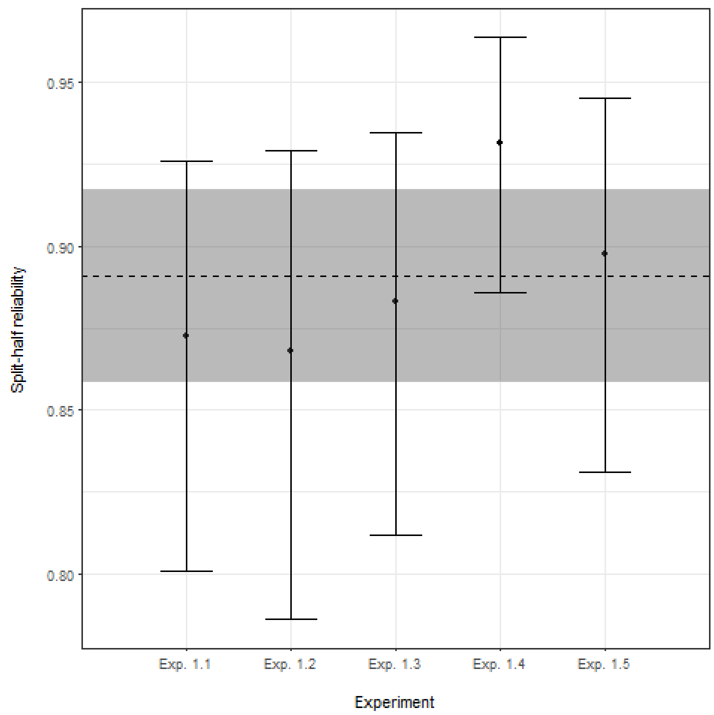

Figure 2. Estimated split-half reliability of Experiments 1.1-1.5 with 95% confidence intervals mapped on the y-axis. The grey shaded areas depict the 95% confidence interval of the weighted mean of the reliability estimates.

Experiment 2: Within Session Repeated Measures Reliability

Although the split-half analysis provides a good estimate of reliability, there are some challenges in the present case. First, different words are used across the two samples for each participant in each simulated set. Second, the participant is in the same experimental session and thus the same mental state—if we want to conclude that this task provides a stable measure of valence we need to test participants multiple times. Specifically, we need to assess the ability of the task to elicit similar responses to multiple presentations of the same stimuli. The present experiment used a test–retest paradigm where the same words were presented twice within a one-hour experimental session. Both blocks, which were completed sequentially without any substantial break, contained identical stimuli presented in pseudorandom order. In order to test the reliability of distance as a metric of valence, we evaluated the correlation between two slopes (one from each block) of the best-fitting regression line where valence predicts distance from the word (Analysis 1). Unlike Experiments 1.1-1.5, we used all words in the two blocks to compute the regression slope estimates. In order to test the reliability of distance to a word as a behavioral outcome (regardless of valence), we evaluated the correlation between responses to any given word produced by each participant in block 1 vs block 2 (Analysis 2).

Method

Participants. Sixty-four undergraduate students at McMaster University in Hamilton, ON participated in this experiment for partial course credit. The data from thirteen participants were removed (seven participants did not make a response on more than 25% of the trials (38 trials), and an additional five participants were not native speakers of English). The remaining 52 (43 women, 48 right-handed) participants ranged in age from 17 to 22 (M = 18.75, SD = 1.17).

Affective Stimuli. We randomly selected half of the words (125 words) from Warriner et al. (2017) Experiment 1: for word selection criteria see Experiment 1. All participants in this experiment saw the same 125 words.

Procedure. Participants were tested in groups of ten or less in a computer lab. Each participant was seated in front of a monitor with a screen resolution of 1024 x 768. After completing a set of demographic questions (including age, sex, handedness and education), participants were instructed to complete the slider task. The task began with a fixation cross centred on the screen; each trial started by clicking on the cross. The affixation cross was replaced with a humanoid manikin centred on a vertical line in the center of the computer screen; the word for that trial was randomly presented at either the top or the bottom of the vertical line, and participants were instructed to slide the manikin (or click a destination on the line) however close or far they wanted the manikin to be from the word. The instructions were the same as the present Experiments 1.1-1.4. After the participants moved the manikin to its final position, they clicked a Submit button, located to the right of the slider (see Figure 1). The experiment was programmed using Experiment Builder software (SR Research, Kanata, ON, Canada).

Participants completed five practice trials, and then asked if they had any questions before the experiment proceeded. Each participant saw each of the 125 words twice, once in the first half of the Experiment (Block 1), and once in the second half (Block 2). The word order was randomized separately in both blocks of the Experiment. The participants were not made aware of the two separate blocks, and were not told the test–retest purpose of the study. Because of a programming error, some participants saw an additional five words for a third time at the end of the experiment. These additional responses were removed from data analysis.

Results and Discussion

Analysis 1 The variables used in this analysis were the same as those used in Experiment 1 with the manikin's distance from the word as the dependent variable and word valence as the independent variable.

To compare individual performance between experimental blocks, we first estimated the effect of valence on the distance of the manikin from the word for each participant and for each half of the experiment. Specifically, for each participant we fitted two ordinary linear regression models (one for 125 words in each half) estimating the effect of valence on the manikin's distance from the word. We operationalized the effect as the slope (the beta coefficient) of the regression line. Since words presented in both blocks were identical (though ordered differently), a reliable performance would elicit similar slopes in an individual, throughout the experiment. This step generated two sets of 52 slope coefficients (corresponding to 52 participants and the two blocks). We then calculated the correlation between the two sets of regression coefficients as a measure of the test–retest reliability within one experimental session.

The distance of the manikin from the word was negatively correlated with valence of the word in both blocks (first block: Pearson's r = -0.572, df = 6498, 95% CI [-0.587,-0.555], $p < .001$; second block: Pearson's r = -0.570, df = 6498, 95% CI [-0.586,-0.553], $p < .001$) as well as the whole experiment together (Pearson's r = -0.570, df = 12998, 95% CI [-0.582,-0.559], $p < .001$). As observed in our previous work (Experiment 1; Warriner et al., 2017), participants moved the manikin closer to positive words, and moved it further away from negative words.

The correlation between two sets of slope coefficients was high (r = 0.804, df = 50, 95% CI [0.680, 0.883], $p < .001$). This demonstrates that valence affects participants' performance in the slider task similarly in the first and second half of the experiment, and also elicits similar responses to the same stimuli presented a second time. The slider task has a high test-retest reliability within one experimental session.

Analysis 2 The variables used in this analysis are the manikin's distance from the word as the dependent variable and block number as the independent variable.

We compared the distance from the word within participant across the two blocks. Specifically, for each word, we generated two lists of distance values for block 1 and block 2 respectively: e.g., distance to word "chore" produced by participant 1 in block 1 would make the first element in list 1, and distance to the same word produced by the same participant in block 2 would make the first element in list 2. Thus, each list contained 52 elements for each participant, repeated for each word. We correlated the two lists for each word, with the correlation coefficient as a measure of how reliably that word is rated within participant across blocks. The second step was to evaluate the central tendency and dispersion of correlation strength across words. The mean correlation, or the reliability of the distance measurement, within a single session was moderately high (r_{mean} = 0.603, SD = 0.155, CI [0.297, 0.867]). This demonstrates that the participants reliably responded with a similar distance in the first and second block within a single session.

<u>Experiment 3: Across Session Test Retest Reliability</u>

In Experiment 2, we showed that participants' performance was replicable within one test–retest session. Another aspect of reliability is whether the task elicits similar responses to the same stimuli over time. We presented the same slider task to participants twice, with experimental sessions separated by a week. The task was expected to produce similar results over the two sessions.

Method

Participants. Forty-five undergraduate students at McMaster University participated in this study for partial course credit. None of the participants had participated in similar studies prior to completing this study. Forty-two of the participants attended the second retest part of the experiment. The remaining three participants were excluded from data analysis. Three additional participants were removed from analysis because they did not respond to more than 25% of trials (62 trials). An additional three participants were removed from analysis due to not being native speakers of English. The remaining 36 native English speakers (33 female) ranged in age from 18 to 27 (M = 18.69, SD =1.58).

Affective Stimuli. The stimuli used in this experiment is the same list of 250 words from Warriner et al. (2017), as described in Experiment 1.

Procedure. The procedure for this experiment was nearly identical to Experiment 2. Participants signed up to take part in a two-part study. The second session took place a week after the first session. The instructions given to participants were the exact same as those given in Experiment 2. Participants were first shown 5 practice words, and then were presented with 250 experimental trials. The 250 words were presented in one block, and were randomly ordered. The participants were given the same set of instructions and the same words the second time they came in for the study, though the order of words in each session was randomized differently. They were not told that the words presented would be the same as the first session, and they were not informed of the purpose of the study. Because of a programming error, some participants saw some words for a third or fourth time at the end of the experiment. These additional responses were removed from data analysis.

Results and Discussion

Analysis 1: The variables used in this analysis were the same as those used in Experiment 1 and Analysis 1 of Experiment 2, with the manikin's distance from the word as the dependent variable and word valence as the independent variable.

To measure the reliability of individual performance in the slider task over time, we first estimated the effect of valence on the distance of the manikin from the word for each participant and for each half of the experiment, using ordinary linear regression models (see Experiment 2). This step generated two sets of 36 slope coefficients (corresponding to 36 participants and the two blocks). We then calculated the correlation between the two sets of regression coefficients as a measure of the test-retest reliability within two experimental sessions separated by one week.

The distance of the manikin was negatively correlated with the valence of the word for both sessions (First session: Pearson's r = -0.559, df = 9178, 95% CI [-0.573,-0.544], p < .001; second session: Pearson's r = -0.532, df = 9178, 95% CI [-0.546,-0.517], p < .001) and the two sessions together (Pearson's r = -0.545, df = 18358, 95% CI [-0.555,-0.535], p < .001). Participants moved the manikin closer to positive words, and moved it further away from negative words.

We calculated the effect of valence on distance for each participant and for each experimental session. A correlation between the two sets of slope coefficients was strong and significant (Pearson's r = 0.754, df = 34, 95% CI [0.566, 0.868], p < .001). This demonstrates that valence of the same set of stimuli affects participants' performance in the slider task similarly in the first and second experimental sessions. The slider task has high test-retest reliability over time.

Analysis 2: The variables and statistical analysis is identical to Analysis 2 of Experiment 2; the manikin's distance from the word as the dependent variable and block number as the independent variable. The only exception is the difference between block 1 and block 2 in Experiment 2 are now represented as session 1 and session 2.

The mean correlation, or the reliability of the distance measurement, across two experimental sessions was moderate (r_{mean} = 0.436, SD = 0.201, CI [0.047, 0.805]). This demonstrates that the participants reliably responded with a similar distance in the first and second experimental session.

The reliability in Experiment 3 is, unsurprisingly, lower than in Experiment 2 (r = 0.754 & 0.804, respectively). This is due to the difference in the length of time between the retest sessions; Experiment 2 had no time between sessions, whereas Experiment 3 had a week between sessions. The same decrease is found in the reliability of the distance measurement, where a separation by a week (r_{mean} = 0.436) caused more variability in responses than the same-day session ((r_{mean} = 0.603).

General Discussion

The slider task provides a reliable, valid and easily reproducible way to capture human affective judgements of words presented in isolation. We hypothesize that this conclusion can be extended to include other types of stimuli, such as pictures, sounds or linguistic phrases. As argued in Warriner et al. (2017), the slider task has addressed numerous criticisms relating to the validity of rating-scale affective judgements. The slider task removes any emotional anchors from the instructions that may potentially affect participants' ratings (see Westbury et al., 2015). Warriner et al. (2017) demonstrated that the functional relationship between valence and distance was not affected by the presence or absence of anchors. The present study additionally demonstrates that there is no difference in the reliability of the task (see Experiments 1.1-1.4 with anchors, and Experiment 1.5 without anchors). Critically, the slider task also allows for a more fine-grained measure of affective judgements, and robust detection of subtle individual differences in response patterns on an interval scale. A question that was left unanswered in Warriner et al. (2017) is whether the slider task and its measurement of individual differences is reliable.

This paper confirms that the slider task performed well on many psychometric reliability measures. In Experiment 1, we demonstrated that the slider task has a high split-half reliability across multiple studies. Participants responded to a words' valence with a high degree of similarity across randomly divided halves of experiments. In Experiment 2, we further confirmed that the slider task is reliable across one hour-long session where a set of words were presented twice. Participants responded similarly to a word's valence in both the first and second half of the session. In Experiment 3, we showed that the slider task is reliable over time. Participants were affected by valence similarly within two sessions which presented the same words that were separated by a week. We also observed in Analyses 2 of Experiments 2 and 3 that the reliability of producing a distance value is consistently lower than the reliability of that distance value as an index of valence. In other words, over two sessions participants were less consistent in how they approached or withdrew from a given word, but in both sessions those responses were driven by the word's valence to a greater degree and with greater consistency. This suggests that absolute values of the distance to words varying in valence are less reliable (and theoretically less informative) than the values of distance relative to the word's valence.

We believe these reliability results, along with the findings from Warriner et al. (2017), show that a sliding scale is a useful tool for collecting affective responses to words, in that it is both valid and reliable. The slider scale shows added utility over a typical discrete ratings scale. Both scales produce similar aggregate data, but the continuous slider scale allows for more detailed responses that are crucial when studying individual differences. As well, unlike a ratings scale, a sliding scale does not rely on anchor terms that are related to valence, a problem which was discussed in Westbury, Keith, Briesemeister, Hofmann, and Jacobs (2015). Additionally, affective ratings are often used in analyses as an interval data, even though they collect ordinal data. The data obtained from the sliding scale are interval.

One possible limitation of the slider task is discussed at length in Warriner et al. (2017): the slider task may be tapping into valence or a related but different psychological construct such as approach-avoidance behaviour, or both. Valence is known to be strongly linked to motivational systems which drive appetitive and aversive responses, to a degree that disentangling these constructs presents a considerable difficulty (see among others, Carver & White, 1994; Lang, 1995). The present method does not enable us to partial out an independent contribution of motivational systems to individual affective responses. While the slider scale is demonstrably a reliable tool for quantifying valence, the question of its link to the approach-avoidance behavior will need to be studied further.

Moreover, it is important to note that we are looking at only one of the two dimensions of affect (J. Russell, 1980), i.e., valence and not arousal. Valence is the measurement of negativity to positivity, whereas arousal is the measurement of calmness to excitedness (J. Russell, 1980). The wording of our sliding scale only taps into valence measurement, whereas other slider scales were able to capture arousal as well (Betella & Verschure, 2016).

In sum, multiple methods confirm that the slider task is very reliable within participants. Additionally, our split-half analysis of a combined dataset (see Exp. 1.6) showed the task's reliability between participants. We conclude that the task has the high degree of reliability that is essential not only for the overall utility of the task for measuring affective responses, but also for its ability to detect subtle patterns of individual variability.

Acknowledgments

This work was supported by the Patterson-Wilson Ontario Graduate Scholarship (OGS) to Constance Imbault. Victor Kuperman's contribution was supported in part by the Natural Sciences and Engineering Research Council or Canada (NSERC) Discovery Grant 402395–2012, the NIH R01 HD 073288 (PI Julie A. Van Dyke), the SSHRC Partnership Training grant 895-2016-1008 (PI Gary Libben), the Early Research Award from the Ontario Ministry of Research and Innovation, and the Canada Research Chair (Tier 2) award. David I. Shore was supported by an NSERC Discovery Grant.

References

Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review*, *96*(2), 358–372. https://doi.org/10.1037/0033-295X.96.2.358

Albaum, G., Best, R., & Hawkins, D. E. L. (1981). Continuous vs Discrete Semantic Differential Rating Scales. *Psychological Reports*, *49*, 83–86.

Allen, N. B., Trinder, J., & Brennan, C. (1999). Affective Startle Modulation in Clinical Depression: Preliminary Findings. *Biological Psychiatry*, *46*(4), 542–550. https://doi.org/https://doi.org/10.1016/S0006-3223(99)00025-6

Altarriba, J., & Bauer, L. M. (2004). The Distinctiveness of Emotion Concepts: A Comparison between Emotion, Abstract, and Concrete Words. *The American Journal of Psychology*, *117*(3), 389–410.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (DSM-5®). *American Psychiatric Pub*.

Anooshian, L. J., & Hertel, P. T. (1994). Emotionality in free recall: Language specificity in bilingual memory. *Cognition and Emotion*, *8*(6), 503–514. https://doi.org/10.1080/02699939408408956

Ayçiçeği-Dinn, A., & Caldwell-Harris, C. L. (2009). Emotion-memory effects in bilingual speakers: A levels-of-processing approach*. *Bilingualism: Language and Cognition*, *12*(03), 291. https://doi.org/10.1017/S1366728909990125

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(04), 577–660. https://doi.org/10.1017/S0140525X99002149

Batson, C. D., Early, S., & Salvarani, G. (1997). Perspective Taking: Imagining How Another Feels Versus Imaging How You Would Feel. *Personality and Social Psychology Bulletin*.

https://doi.org/10.1177/0146167297237008

Bayer, M., Sommer, W., & Schacht, A. (2011). Emotional words impact the mind but not the body: Evidence from pupillary responses. *Psychophysiology*, *48*(11), 1554–1562. https://doi.org/10.1111/j.1469-8986.2011.01219.x

Betella, A., & Verschure, P. F. M. J. (2016). The affective slider: A digital self-assessment scale for the measurement of human emotions. *PLoS ONE*, *11*(2), 1–11. https://doi.org/10.1371/journal.pone.0148037

Bonin, P., Aubert, L., Malardier, N., & Niedenthal, P. M. (2016). Subjective Et De Valence Émotionnelle Pour 866 Mots, *103*, 655–694.

Bradley, M. M., & Lang, P. P. J. (1999). Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings. *Psychology*, *Technical*(C–1), 0. https://doi.org/10.1109/MIC.2008.114

Brase, J., & Mani, N. (2017). Effects of learning context on the acquisition and processing of emotional words in bilinguals. *Emotion*, *17*(4), 628–639. https://doi.org/10.1037/emo0000263

Briesemeister, B. B., Kuchinke, L., & Jacobs, A. M. (2011). Discrete emotion effects on lexical decision response times. *PLoS ONE*, *6*(8). https://doi.org/10.1371/journal.pone.0023743

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990. https://doi.org/10.3758/BRM.41.4.977

Bylsma, L. M., Morris, B. H., & Rottenberg, J. (2008). A meta-analysis of emotional reactivity in major depressive disorder. *Clinical Psychology Review*, *28*(4), 676–691. https://doi.org/10.1016/j.cpr.2007.10.001

Caldwell-Harris, C. L. (2015). Emotionality Differences Between a Native and Foreign Language. *Current Directions in Psychological Science*, *24*(3), 214–219. https://doi.org/10.1177/0963721414566268

Caldwell-Harris, C. L., Tong, J., Lung, W., & Poo, S. (2011). Physiological reactivity to emotional phrases in Mandarin--English bilinguals. *International Journal of Bilingualism*, *15*(3), 329–352. https://doi.org/10.1177/1367006910379262

Canli, T., Sivers, H., Whitfield, S. L., Gotlib, I. H., & Gabrieli, J. D. E. (2002). Amygdala response to happy faces as a function of extraversion. *Science*, *296*(5576), 2191. https://doi.org/10.1126/science.1068749

Cartwright, C., Gibson, K., Read, J., Cowan, O., & Dehar, T. (2016). Long-term antidepressant use: patient perspectives of benefits and adverse effects. *Patient Preference and Adherence*, *Volume 10*, 1401–1407. https://doi.org/10.2147/PPA.S110632

Carver, C. S., & White, T. L. (1994). Behavioral Inhibition, Behavioral Activation and affective responses to impending reward and punishment. *Journal of Personality and Social Psychology*, *67*, 319–333.

Chen, M., & Bargh, J. A. (1997). Nonconscious behavioral confirmation processes: The self-fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology*, *33*(5), 541–560. https://doi.org/10.1006/jesp.1997.1329

Citron, F. M. M., Weekes, B. S., & Ferstl, E. C. (2013). Effects of valence and arousal on written word recognition: Time course and ERP correlates. *Neuroscience Letters*, *533*, 90–95.

Clark, J., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*(3), 149–210. https://doi.org/10.1007/BF01320076

Cutrona, C. E., & Troutman, B. R. (1986). Social Support , Infant Temperament , and Parenting Self-Efficacy : A Mediational Model of Postpartum Depression Author (s): Carolyn E . Cutrona and Beth R . Troutman Published by : Wiley on behalf of the Society for Research in Child Development Stable, *57*(6), 1507–1518.

Daneman, M., & Carpenter, P. A. (1980). Individual Differences in Working Memory and Reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466.

De Grauwe, S., Willems, R. M., Rueschemeyer, S. A., Lemhöfer, K., & Schriefers, H. (2014). Embodied language in first- and second-language speakers: Neural correlates of processing motor verbs. *Neuropsychologia*, *56*(1), 334–349.

https://doi.org/10.1016/j.neuropsychologia.2014.02.003

De Pascalis, V., Fiore, A. D., & Sparita, A. (1996). Personality, event-related potential (ERP) and heart rate (HR): An investigation of Gray's theory. *Personality and Individual Differences*, *20*(6), 733–746. https://doi.org/10.1016/0191-8869(96)00016-5

De Pascalis, V., Strippoli, E., Riccardi, P., & Vergari, F. (2004). Personality, event-related potential (ERP) and heart rate (HR) in emotional word processing. *Personality and Individual Differences*, *36*(4), 873–891. https://doi.org/10.1016/S0191-8869(03)00159-4

Decety, J., & Batson, C. D. (2007). Social neuroscience approaches to interpersonal sensitivity. *Social Neuroscience*, *2*(3–4), 151–157. https://doi.org/10.1080/17470910701506060

Degner, J., Doycheva, C., & Wentura, D. (2012). It matters how much you talk: On the automaticity of affective connotations of first and second language words. *Bilingualism*, *15*(1), 181–189. https://doi.org/10.1017/S1366728911000095

Dewaele, J.-M. (2004). The Emotional Force of Swearwords and Taboo Words in the Speech of Multilinguals. *Journal of Multilingual and Multicultural Development*, *25*(23), 204–222. https://doi.org/10.1080/01434630408666529

Dewaele, J. M., & Nakano, S. (2013). Multilinguals' perceptions of feeling different when switching languages. *Journal of Multilingual and Multicultural Development*, *34*(2), 107–120. https://doi.org/10.1080/01434632.2012.712133

Dresler, T., Mériau, K., Heekeren, H. R., & Van Der Meer, E. (2009). Emotional Stroop task: Effect of word arousal and subject anxiety on emotional interference. *Psychological Research*, *73*(3), 364–371. https://doi.org/10.1007/s00426-008-0154-6

Ehrman, M. E., Leaver, B. Lou, & Oxford, R. L. (2003). A brief overview of individual differences in second language learning. *System*, *31*(3), 313–330. https://doi.org/10.1016/S0346-251X(03)00045-9

Eilola, T. M., Havelka, J., & Sharma, D. (2007). Emotional activation in the first and second language. *Cognition and Emotion*, *21*(5), 1064–1076. https://doi.org/10.1080/02699930601054109

Eisman, A. B., Stoddard, S. a, Heinze, J., Caldwell, C. H., & Zimmerman, M. a. (2015). Depressive symptoms, social support, and violence exposure among urban youth: A longitudinal study of resilience. *Developmental Psychology*, *51*(9), 1307–1316. https://doi.org/10.1037/a0039501

Ekman, P. (1992). An Argument for Basic Emotions. *Cognition and Emotion, 6*(3–4), 169–200. https://doi.org/10.1080/02699939208411068

Elliot, A. J. (2006). The Hierarchical Model of Approach-Avoidance Motivation. *Motivation and Emotion, 30*(2), 111–116. https://doi.org/10.1007/s11031-006-9028-7

Epstein, J., Pan, H., Kocsis, J., Yang, Y., Butler, T., Chusid, J., … Silbersweig, D. (2006). Lack of Ventral Striatal Response to Positive Stimuli inDepressed Versus Normal Subjects. *The American Journal of Psychiatry*, 1784. https://doi.org/10.1176/ajp.2006.163.10.1784

Farrow, T. F. D., & Woodruff, P. W. R. (2008). Empathy in Mental Illness. *Psychiatric Services*, *59*, 581–581. https://doi.org/10.1176/appi.ps.59.5.581

Feldman Barrett, L., & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology*, *74*(4), 967–984. https://doi.org/10.1037/0022-3514.74.4.967

Ferré, P., García, T., Fraga, I., Sánchez-Casas, R., & Molero, M. (2010). Memory for emotional words in bilinguals: Do words have the same emotional intensity in the first and in the second language? *Cognition and Emotion*, *24*(5), 760–785. https://doi.org/10.1080/02699930902985779

Foroni, F. (2015). Do we embody second language? Evidence for ‘partial’ simulation during processing of a second language. *Brain and Cognition*, *99*, 8–16. https://doi.org/10.1016/j.bandc.2015.06.006

George, L., Blazer, D., Hughes, D., & Fowler, N. (1989). Social support and the outcome of major depression. *British Journal of Psychiatry*, *154*, 478–485.

Golin, S., Hartman, S. A., Klatt, E. N., Munz, K., & Wolfgang, G. L. (1977). Effects of Self-esteem Manipulation on Arousal and Reactions to Sad Models in Depressed and Nondepressed College Students. *Journal of Abnormal Psychology*, *86*(4), 435–439.

Haaga, D. A. F., & Beck, A. T. (1995). Perspectives on depressive realism: Implications for cognitive theory of depression. *Behaviour Research and Therapy*, *33*(1), 41–48. https://doi.org/10.1016/0005-7967(94)E0016-C

Harris, C. L., Ayçiçeği, A., & Gleason, J. B. (2003). Taboo words and reprimands elicit greater autonomic reactivity in a first language than in a second language. *Applied Psycholinguistics*, *24*(4), 561–579. https://doi.org/10.1017/S0142716403000286

Harris, C. L., Gleason, J. B., & Ayçiçeği, A. (2006). When is a First Language More Emotional? Psychophysiological Evidence from Bilingual Speakers. In *Bilingual minds: Emotional experience, expression and representation* (pp. 257–283). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.69.5830%0Ahttp://www.zeit.de/wirtschaft/2014-08/russland-import-sanktionen%0Ahttp://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.69.5830%0Ahttp://www.zeit.de/wirtschaft/2014-08/russland-import-sa

Hellman, A. B. (2011). Vocabulary size and depth of word knowledge in adult-onset second language acquisition. *International Journal of Applied Linguistics*, *21*(2), 162–182. https://doi.org/10.1111/j.1473-4192.2010.00265.x

Hinojosa, J. a., Martínez-García, N., Villalba-García, C., Fernández-Folgueiras, U., Sánchez-Carmona, a., Pozo, M. a., & Montoro, P. R. (2015). Affective norms of 875 Spanish words for five discrete emotional categories and two emotional dimensions. *Behavior Research Methods*, (August). https://doi.org/10.3758/s13428-015-0572-5

Imbault, C., Shore, D., & Kuperman, V. (2017). Reliability of the sliding scale for collecting affective responses to words. *Behavior Research Methods*.

Inoue, Y., Yamada, K., & Kanba, S. (2006). Deficit in theory of mind is a risk for relapse of major depression. *Journal of Affective Disorders*, *95*(1–3), 125–127. https://doi.org/10.1016/j.jad.2006.04.018

Inoue, Yumiko, Tonooka, Y., Yamada, K., & Kanba, S. (2004). Deficiency of theory of mind in patients with remitted mood disorder. *Journal of Affective Disorders*, *82*(3), 403–409. https://doi.org/10.1016/j.jad.2004.04.004

Itkes, O., Eviatar, Z., & Kron, A. (2019). Semantic and affective manifestations of ambi (valence). *Cognition and Emotion*, *9931*. https://doi.org/10.1080/02699931.2018.1564249

Itkes, O., Kimchi, R., Kron, A., & Carmel, M. (2017). Dissociating Affective and Semantic Valence, *146*(7), 1–36. https://doi.org/10.1037/xge0000291

Itzhak, I., Vingron, N., Baum, S. R., & Titone, D. (2017). Bilingualism in the real world: How proficiency, emotion, and personality in a second language impact communication in clinical and legal settings. *Translational Issues in Psychological Science*, *3*(1), 48–65. https://doi.org/10.1037/tps0000103

Johnson-Laird, P. N., & Oatley, K. (1989). The Language of Emotions: An Analysis of a Semantic Field. *Cognition and Emotion*, *3*(2), 81–123. https://doi.org/10.1080/02699938908408075

Jolliffe, D., & Farrington, D. P. (2006). Development and validation of the Basic Empathy Scale. *Journal of Adolescence*, *29*(4), 589–611. https://doi.org/10.1016/j.adolescence.2005.08.010

Just, M. A., & Carpenter, P. A. (1999). A Capacity Theory of Comprehension : Indicidual Dirrefences in Working Memory, (January), 1–16.

Kazanas, S. A., & Altarriba, J. (2016). Emotion Word Processing: Effects of Word Type and Valence in Spanish–English Bilinguals. *Journal of Psycholinguistic Research*, *45*(2), 395–406. https://doi.org/10.1007/s10936-015-9357-3

Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual Differences in Language Acquisition and Processing. *Trends in Cognitive Sciences*, *22*(2), 154–169. https://doi.org/10.1016/j.tics.2017.11.006

Kleinginna, P. R., & Kleinginna, A. M. (1981). A categorized list of motivation definitions, with a suggestion for a consensual definition. *Motivation and Emotion*, *5*(3), 263–291. https://doi.org/10.1007/BF00993889

Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The Representation of Abstract Words: Why Emotion Matters. *Journal of Experimental Psychology: General*, *140*(1), 14–34. https://doi.org/10.1037/a0021446

Kousta, S. T., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition*, *112*(3), 473–481. https://doi.org/10.1016/j.cognition.2009.06.007

Kuchinke, L., Võ, M. L. H., Hofmann, M., & Jacobs, A. M. (2007). Pupillary responses during lexical decisions vary with word frequency but not emotional valence. *International Journal of Psychophysiology*, *65*(2), 132–140. https://doi.org/10.1016/j.ijpsycho.2007.04.004

Kunzmann, U., & Grühn, D. (2005). Age differences in emotional reactivity: The sample case of sadness. *Psychology and Aging*, *20*(1), 47–59. https://doi.org/10.1037/0882-7974.20.1.47

Kuperman, V., Estes, Z., Brysbaert, M., & Warriner, A. B. (2014). Emotion and language: Valence and arousal affect word recognition. *Journal of Experimental …*, (4). Retrieved from http://psycnet.apa.org/psycarticles/2014-03909-001

Kuperman, V., & Van Dyke, J. a. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language*, *65*(1), 42–73. https://doi.org/10.1016/j.jml.2011.03.002

Kuppens, P., Tuerlinckx, F., Yik, M., Koval, P., Coosemans, J., Zeng, K. J., & Russell, J. A. (2017). The Relation Between Valence and Arousal in Subjective Experience Varies With Personality and Culture. *Journal of Personality*, *85*(4), 530–542. https://doi.org/10.1111/jopy.12258

Kuznetsova, A., Brockhoff, P., & Christensen, R. (2016). lmerTest: Tests in Linear Mixed Effects Models. *R Package Version*, *3.0.0,* https://cran.r-project.org/package=lmerTest. Retrieved from https://cran.r-project.org/package=lmerTest

Lang, P. J. (1995). The Emotion Probe: Studies of Motivation and Attention. *American Psychologist*. https://doi.org/10.1037/0003-066X.50.5.372

Larsen, R. J., & Ketelaar, T. (1991). Personality and susceptibility to positive and negative emotional states. *Journal of Personality and Social Psychology*, *61*(1), 132–140. https://doi.org/10.1037/0022-3514.61.1.132

Levenson, R. W., Carstensen, L. L., Friesen, W. V., & Ekman, P. (1991). Emotion, physiology, and expression in old age. *Psychology and Aging*, *6*(1), 28–35.

https://doi.org/10.1037/0882-7974.6.1.28

Lewinsohn, P. M., Lobitz, W. C., & Wilson, S. (1973). "Sensitivity" of Depressed Individuals to Aversive Stimuli. *Journal of Abnormal Psychology*, *81*(3), 259–263.

Long, L. S., & Titone, D. A. (2007). Psychopathy and verbal emotion processing in non-incarcerated males. *Cognition and Emotion*, *21*(1), 119–145. https://doi.org/10.1080/02699930600551766

Mackay, D. G., Shafto, M., Taylor, J. K., Marian, D. E., Abrams, L., & Dyer, J. R. (2004). Relations between emotion, memory, and attention: Evidence from taboo Stroop, lexical decision, and immediate memory tasks. *Memory & Cognition*, *32*(3), 474–488. https://doi.org/10.3758/BF03195840

Macleod, J. W., Lawrence, M. a, McConnell, M. M., Eskes, G. a, Klein, R. M., & Shore, D. I. (2010). Appraising the ANT: Psychometric and theoretical considerations of the Attention Network Test. *Neuropsychology*, *24*(5), 637–651. https://doi.org/10.1037/a0019803

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research : JSLHR*, *50*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

Mcwhirter, B. T., Besett-Alesch, T. M., Horibata, J., & Gat, I. (2002). Loneliness in High Risk Adolescents: The Role of Coping, Self-Esteem, and Empathy. *Journal of Youth Studies*, *5*(1), 69–84. https://doi.org/10.1080/13676260120111177

Méndez-Bértolo, C., Pozo, M. A., & Hinojosa, J. A. (2011). Word frequency modulates the processing of emotional words: Convergent behavioral and electrophysiological data. *Neuroscience Letters*, *494*(3), 250–254. https://doi.org/10.1016/j.neulet.2011.03.026

Monnier, C., & Syssau, A. (2014). Affective Norms for French Words (FAN). *Behavior Research Methods*, *46*(4), 1128–1137. https://doi.org/10.3758/s13428-013-0431-1

Moors, A., De Houwer, J., Hermans, D., Wanmaker, S., van Schie, K., Van Harmelen, A.-L., … Brysbaert, M. (2013). Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words. *Behavior Research Methods*, *45*(1), 169–177.

https://doi.org/10.3758/s13428-012-0243-8

Mor, N., & Winquist, J. (2002). Self-focused attention and negative affect: A meta-analysis. *Psychological Bulletin*, *128*(4), 638–662. https://doi.org/10.1037/0033-2909.128.4.638

O'Connor, L. E., Berry, J. W., Lewis, T., Mulherin, K., & Crisotomo, P. S. (2007). Empathy and depression: the moral system on overdrive. *Empathy in Mental Ilness*. https://doi.org/10.1017/CBO9780511543753.005

Orne, M. T. (1962). On the Social Psychology of Psychological Experiment: With Particular Reference to Demand Characteristics and Their Implications. *American Psychologist*, *17*(11), 776–783. Retrieved from https://journals.scholarsportal.info/details/0003066x/v17i0011/776_otspottdcati.xml

Ostaszewski, K., & Zimmerman, M. A. (2006). The effects of cumulative risks and promotive factors on urban adolescent alcohol and other drug use: A longitudinal study of resiliency. *American Journal of Community Psychology*, *38*(3–4), 237–249. https://doi.org/10.1007/s10464-006-9076-x

Pearson, C., Janz, T., & Ali, J. (2013). *Mental and substance use disorders in Canada*.

Peirce, R., Frone, M., Russell, M., Cooper, M., & Mudar, P. (2000). A Longitudinal Model of Social Contact, Social Support, Depression, and Alcohol Use. *Health Psychol*, *19*(1), 28–38. https://doi.org/10.1037/0278-6133.19.1.28

Pennebaker, J., Booth, R., & Francis, M. (2007). Linguistic inquiry and word count: LIWC. Austin, TX: liwc net.

Perfetti, C. (2007). Reading Ability: Lexical Quality to Comprehension. *Scientific Studies of Reading*, *11*(4), 357–383. https://doi.org/10.1080/10888430701530730

Ponari, M., Rodriguez-Cuadrado, S., & Vinson, D. (2014). Processing advantage for emotional words in bilingual speakers. *Emotion*, *15*(January), 644–652. https://doi.org/10.1037/emo0000061

Ponari, M., Rodriguez-Cuadrado, S., Vinson, D., Fox, N., Costa, A., & Vigliocco, G. (2014). Processing advantage for emotional words in bilingual speakers. *Emotion*, *15*(January), 644–652. https://doi.org/10.1037/emo0000061

Pratto, F., & John, O. P. (1991). Automatic Vigilance: The Attention-Grabbing Power of Negative Social Information. *Journal of Personality and Social Psychology*, *61*(3), 380–391. https://doi.org/10.1037/0022-3514.61.3.380

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *The Behabioral and Brain Sciences*, *4*, 515–526. https://doi.org/10.1016/j.celrep.2011.1011.1001.7.

Puntoni, S., de Langhe, B., & van Osselaer, S. M. J. (2009). Bilingualism and the Emotional Intensity of Advertising Language. *Journal of Consumer Research*, *35*(6), 1012–1025. https://doi.org/10.1086/595022

Pyszczynski, T., Holt, K., & Greenberg, J. (1987). Depression, self-focused attention, and expectancies for positive and negative future life events for self and others. *Journal of Personality and Social Psychology*, *52*, 994–1001. https://doi.org/10.1037/0022-3514.52.5.994

R Core Team, I. (2017). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna (2014). R Foundation for Statistical Computing Authors R Core Team.

Recio, G., Conrad, M., Hansen, L. B., & Jacobs, A. M. (2014). On pleasure and thrill: The interplay between arousal and valence during visual word recognition. *Brain and Language*, *134*, 34–43. https://doi.org/10.1016/j.bandl.2014.03.009

Redondo, J., Fraga, I., Padrón, I., & Comesaña, M. (2007). The Spanish adaptation of ANEW (affective norms for English words). *Behavior Research Methods*, *39*(3), 600–605. https://doi.org/10.3758/BF03193031

Rickard, N. S. (2004). Intense emotional responses to music: a test of the physiological arousal hypothesis. *Psychology of Music*, *32*(4), 371–388. https://doi.org/10.1177/0305735604046096

Rottenberg, J., & Gotlib, I. (2004). Socioemotional Functioning in Depression. In *Mood Disorders: A Handbook of Science and Practice* (pp. 61–77).

Rottenberg, J., Gross, J. J., & Gotlib, I. H. (2005). Emotion context insensitivity in major depressive disorder. *Journal of Abnormal Psychology*, *114*(4), 627–639. https://doi.org/10.1037/0021-843X.114.4.627

Rottenberg, J., & Hindash, A. C. (2015). Emerging evidence for emotion context insensitivity in depression. *Current Opinion in Psychology*, *4*, 1–5. https://doi.org/10.1016/j.copsyc.2014.12.025

Rush, J., Trivedi, M. H., Ibrahim, H. M., Carmody, T. J., Arnow, B., Klein, D. N., … Keller, M. B. (2003). The 16-Item Quick Inventory of Depressive Symptomatology. *Depression*, *54*(5), 573–583. https://doi.org/10.1016/S0006-3223(03)01866-8

Russell, C. J., & Bobko, P. (1992). Moderated Regression Analysis and Likert Scales Too Coarse for Comfort. *Journal of Applied Psychology*, *77*(3), 336–342.

Russell, J. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*. Retrieved from http://psycnet.apa.org/psycinfo/1981-25062-001

Sandell, K., & Bornäs, H. (2017). Functioning Numbness Instead of Feelings as a Direction: Young Adults' Experiences of Antidepressant Use. *Sociology*, *51*(3), 543–558. https://doi.org/10.1177/0038038515591947

Santiago-Rivera, A. L., & Altarriba, J. (2002). The Role of Language in Therapy With the Spanish-English Bilingual Client. *Professional Psychology: Research and Practice*, *33*(1), 30–38. https://doi.org/10.1037/0735-7028.33.1.30

Schutte, N. S., Malouff, J. M., Bobik, C., Coston, T. D., Greeson, C., Jedlicka, C., … Wendorf, G. (2001). Emotional intelligence and interpersonal relations. *Journal of Social Psychology*, *141*(4), 523–536. https://doi.org/10.1080/00224540109600569

Scott, G., O'Donnell, P., Leuthold, H., & Sereno, S. (2009). Early emotion word processing: evidence from event-related potentials. *Biological Psychology*, *80*(1), 95–104. https://doi.org/S0301-0511(08)00073-2 [pii]\r10.1016/j.biopsycho.2008.03.010

Scott, G., O'Donnell, P., & Sereno, S. (2012). Emotion words affect eye fixations during reading. *Journal of Experimental Psychology: Learning Memory and Cognition*, *38*(3), 783–792. https://doi.org/10.1037/a0027209

Segalowitz, N., Trofimovich, P., Gatbonton, E., & Sokolovskaya, A. (2008). Feeling affect in a second language: The role of word recognition automaticity. *The Mental Lexicon*, *3*(1), 47–71. https://doi.org/10.1075/ml.3.1.05seg

Sharot, T. (2011). The optimism bias. *Current Biology*, *21*(23), R941–R945. https://doi.org/10.1016/j.cub.2011.10.030

Sheikh, N. A., & Titone, D. A. (2013). Sensorimotor and linguistic information attenuate emotional word processing benefits: An eye-movement study. *Emotion*, *13*(6), 1107–1121. https://doi.org/10.1037/a0032417

Sheikh, N., & Titone, D. (2015). The embodiment of emotional words in a second language: An eye-movement study. *Cognition & Emotion*, (April), 1–13. https://doi.org/10.1080/02699931.2015.1018144

Shields, M., Gorber, S. C., Janssen, I., & Tremblay, M. S. (2011). Bias in self-reported estimates of obesity in Canadian health surveys: An update on correction equations for adults. *Health Reports*, *22*(3).

Sigmon, S. T., & Nelson-Gray, R. O. (1992). Sensitivity to Aversive Events in Depression: Antecedent, Concomitant, or Consequent? *Journal of Psychopathology and Behavioral Assessment*, *14*(3), 225–246.

Skehan, P. (1991). Individual Differences in Second Language Learning. *Studies in Second Language Acquisition*, *13*(02), 275. https://doi.org/10.1017/S0272263100009979

Sloan, D. M., Strauss, M. E., Quirk, S. W., & Sajatovic, M. (1997). Subjective and expressive emotional responses in depression. *Journal of Affective Disorders*, *46*, 135–141. https://doi.org/https://doi.org/10.1016/S0165-0327(97)00097-9

Sloan, D. M., Strauss, M. E., & Wisner, K. L. (2001). Diminished Response to Pleasant Stimuli by Depressed Women. *Journal of Abnormal Psychology*, *110*(3), 488–493. https://doi.org/I0.1037//0021-843X.110.3.488

Smith, D. P., Hillman, C. H., & Duley, A. R. (2005). Influences of Age on Emotional Reactivity During Picture Processing. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *60*(1), P49–P56. https://doi.org/10.1093/geronb/60.1.P49

Soares, A. P., Comesaña, M., Pinheiro, A. P., Simões, A., & Frade, S. (2012). Behavior and Research Methods Corresponding author : *Behavior Research Methods*, *44*(1), 256–269. https://doi.org/10.3758/s13428-011-0131-7

Stadthagen-González, H., Ferré, P., Pérez-Sánchez, M. A., Imbault, C., & Hinojosa, J. A. (2017). Norms for 10,491 Spanish words for five discrete emotions: Happiness, disgust, anger, fear, and sadness. *Behavior Research Methods*. https://doi.org/10.3758/s13428-017-0962-y

Stadthagen-Gonzalez, H., Imbault, C., Pérez Sánchez, M. A., & Brysbaert, M. (2017). Norms of valence and arousal for 14,031 Spanish words. *Behavior Research Methods*, 1–45. https://doi.org/10.3758/s13428-015-0700-2

Statistics Canada. (2013). *National Household Survey*. *Immigration and Ethnocultural Diversity in Canada* (Vol. 99-010-X20). https://doi.org/99-010-X2011001

Sutton, T. M., Altarriba, J., Gianico, J. L., & Basnight-Brown, D. M. (2007). The automatic access of emotion: Emotional Stroop effects in Spanish-English bilingual speakers. *Cognition and Emotion*, *21*(5), 1077–1090. https://doi.org/10.1080/02699930601054133

Thoma, P., Norra, C., Juckel, G., Suchan, B., & Bellebaum, C. (2015). Performance monitoring and empathy during active and observational learning in patients with major depression. *Biological Psychology*, *109*, 222–231. https://doi.org/10.1016/j.biopsycho.2015.06.002

Tsai, J. L., Knutson, B., & Fung, H. H. (2006). Cultural variation in affect valuation. *Journal of Personality and Social Psychology*, *90*(2), 288–307. https://doi.org/10.1037/0022-3514.90.2.288

Velez-Uribe, I., & Rosselli, M. (2019). The auditory and visual appraisal of emotion-related words in Spanish–English bilinguals. *Bilingualism: Language and Cognition*, *22*(1), 30–46. https://doi.org/10.1017/S1366728917000517

Vinson, D., Ponari, M., & Vigliocco, G. (2014). How does emotional content affect lexical processing? *Cognition & Emotion*, *28*(4), 737–746. https://doi.org/10.1080/02699931.2013.851068

Võ, M. L. H., Jacobs, A. M., Kuchinke, L., Hofmann, M., Conrad, M., Schacht, A., & Hutzler, F. (2008). The coupling of emotion and cognition in the eye: Introducing the pupil old/new

effect. *Psychophysiology*, *45*(1), 130–140. https://doi.org/10.1111/j.1469-8986.2007.00606.x

Wang, Y. guang, Wang, Y. qiang, Chen, S. lin, Zhu, C. yan, & Wang, K. (2008). Theory of mind disability in major depression with or without psychotic symptoms: A componential view. *Psychiatry Research*, *161*(2), 153–161. https://doi.org/10.1016/j.psychres.2007.07.018

Waris, O., Soveri, A., Ahti, M., Hoffing, R. C., Ventus, D., Jaeggi, S. M., … Laine, M. (2017). A latent factor analysis of working memory measures using large-scale data. *Frontiers in Psychology*, *8*(JUN), 1–14. https://doi.org/10.3389/fpsyg.2017.01062

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191–1207. https://doi.org/10.3758/s13428-012-0314-x

Warriner, A. B., Shore, D. I., Schmidt, L. A., Imbault, C. L., & Kuperman, V. (2017). Sliding into happiness: A new tool for measuring affective responses to words. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *71*(1), 71–88. https://doi.org/10.1037/cep0000112

Watson, D., Clark, L. A., & Carey, G. (1988). Positive and negative affectivity and their relation to anxiety and depressive disorders. *Journal of Abnormal Psychology*, *97*(3), 346–353. https://doi.org/10.1037/0021-843X.97.3.346

Weisleder, A., & Fernald, A. (2013). Talking to Children Matters: Early Language Experience Strengthens Processing and Builds Vocabulary. *Psychological Science*, *24*(11), 2143–2152. https://doi.org/10.1177/0956797613488145

Westbury, C., Keith, J., Briesemeister, B. B., Hofmann, M. J., & Jacobs, A. M. (2015). Avoid violence, rioting, and outrage; approach celebration, delight, and strength: Using large text corpora to compute valence, arousal, and the basic emotions. *The Quarterly Journal of Experimental Psychology*, *0*(0), 1–24. https://doi.org/10.1080/17470218.2014.970204

Wickham, H., & Chang, W. (2016). Create elegant data visualisations using the grammar of graphics. https://doi.org/10.1093/bioinformatics/btr406

Williams, J. M., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and

psychopathology. *Psychological Bulletin*, *120*(1), 3–24. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8711015

Winskel, H. (2013). The emotional Stroop task and emotionality rating of negative and neutral words in late Thai-English bilinguals. *International Journal of Psychology*. Taylor & Francis. https://doi.org/10.1080/00207594.2013.793800

Wolkenstein, L., Schönenberg, M., Schirm, E., & Hautzinger, M. (2011). I can see what you feel, but i can't deal with it: Impaired theory of mind in depression. *Journal of Affective Disorders*, *132*(1–2), 104–111. https://doi.org/10.1016/j.jad.2011.02.010

Wolter, B. (2001). COMPARING THE L1 AND L2 MENTAL LEXICON A Depth of Individual Word Knowledge Model. *Studies in Second Language Acquisition*, *23*(1), S0272263101001024. https://doi.org/10.1017/S0272263101001024

Zuroff, D., Koestner, R., Moskowitz, D. S., McBride, C., Marshall, M., & Bagby, M. (2007). Autonomous motivation for therapy: A new common factor in brief treatments for depression. *Psychotherapy Research*, *17*(2), 137–148. https://doi.org/10.1080/10503300600919380

Chapter 4

Abstract

The perspective-taking ability to imagine another person's feelings and thoughts is paramount for successful communication. This study pursued two questions regarding the link between perspective-taking and depressive symptomatology in a task where participants provided responses to words ranging in their positivity. First, we examined in a between-participants experimental manipulation how the presence of depressive symptoms influenced participants' emotional reactivity. Second, we measured within-participants, how their responses change as a function of the perspective they are assigned to take, that of a depressed or a non-depressed person. Our main interest is in the interaction of the two effects: we examine how one's emotional state determines the ability to engender someone else's responses. Our central finding is that depressive symptoms lead to emotional insensitivity, i.e., weaker responses to extremely positive and negative words. Furthermore, depressive symptoms come with a much weaker ability to take a non-depressed perspective. Finally, non-depressed participants demonstrated an excellent ability to mimic the blunt affect of depression when responding for the other group, suggesting that the outlook of a depressed individual is available to people throughout the range of depressive symptomatology. We discuss the implications of these findings for quantifying emotional reactivity during depression, as well as the diagnosis and prognosis of depression.

Introduction

The ability to take the perspective of another person, i.e., to imagine that person's feelings and thoughts, is at the core of human social interaction and is the glue for a functional society (Decety & Batson, 2007; Premack & Woodruff, 1978). A central question we ask in this paper is whether the perspective-taking ability undergoes any transformation in depression. This question requires knowledge of how well-developed perspective-taking is in non-depressed individuals and how the presence of depressive symptoms affects it. Specifically, we examine two novel questions: (i) whether people with depressive symptoms can imagine emotional states of non-depressed people and (ii) how accurately can non-depressed people mimic the emotional states of depression. Our choice of depression for pursuing question (i) has several reasons. First, the very definition of this mood disorder cites "hopelessness, anhedonia, loss of interest" and other markers of affective impairment (American Psychiatric Association, 2013). Thus, one's emotional state in depression tends to differ greatly from a non-depressed one. The literature also reports that depression comes with a weakened Theory of Mind and thus a weaker understanding of others' emotions (Premack & Woodruff, 1978; Wang, Wang, Chen, Zhu, & Wang, 2008; Wolkenstein, Schönenberg, Schirm, & Hautzinger, 2011). This is possibly due to increased self-focused attention (Mor & Winquist, 2002), an obvious obstacle to taking another perspective. Whatever the cause, the deficit in perspective-taking is clinically important, because it has been argued to lead to a feeling of interpersonal isolation and loneliness (Decety & Batson, 2007; Mcwhirter, Besett-Alesch, Horibata, & Gat, 2002), exacerbating depression and social anxiety (Eisman, Stoddard, Heinze, Caldwell, & Zimmerman, 2015; George, Blazer, Hughes, & Fowler, 1989; Ostaszewski & Zimmerman, 2006).

Despite the monumental literature of depression and affect, there has been little to no direct research on perspective-taking in this literature, as opposed to related concepts of empathy and Theory of Mind (Farrow & Woodruff, 2008; Y. Inoue, Yamada, & Kanba, 2006; O'Connor, Berry, Lewis, Mulherin, & Crisotomo, 2007; Wang et al., 2008; Wolkenstein et al., 2011). In this paper, we make a distinction between perspective-taking and Theory of Mind. We define Theory of Mind as the ability to understand that others have their own mental states that are different from our own (Premack & Woodruff, 1978). Perspective-taking, on the other hand, is the ability to put yourself into someone else's shoes; to understand and perceive *how* others feel (Batson et al., 1997). Though these terms are related and often confused, we construe them as different constructs. The keyword search with “perspective taking” and “depression” retrieved only 42 PUBMED citations on September 5th, 2017, none of which directly address this relationship, but see (Thoma, Norra, Juckel, Suchan, & Bellebaum, 2015). Intriguingly, the opposite is also true: little is known about whether a non-depressed person can construe the point of view of a person with depression, i.e., our question (ii). Yet there are indications that this ability is important to the prognosis of depression. Increased perceived social support and understanding are negatively correlated with depression severity (Peirce et al., 2000) and protect against postpartum depression (Cutrona & Troutman, 1986). Thus, our second focus is on whether non-depressed individuals can mimic the emotional state of depression.

Our study examines individuals’ responses to a broad range of emotion-laden linguistic stimuli, while those individuals take on a perspective of a person with and without depression. This task is administered within-subjects, giving insight into how strongly the change in perspective influences the magnitude and polarity of emotional responses. We manipulate the required perspective and select individuals with varying severity of depressive symptoms. Specifically, we quantify how the severity of one’s depressive symptoms (as measured via a clinical self-report assessment) modulates one’s ability to take another perspective. There are at least two ways to approach our questions. One possibility is to ask participants to self-identify as depressed or non-depressed, and then instruct them to complete the task for their own group and the opposite group. This approach would afford an insight into a person’s own emotional reactivity, but it would also come with a cost. Specifically, the experimental demand of categorical self-identification can elicit cognitive and social response biases (including the in- and out-group stereotyping) (Orne, 1962) when taking their own purported perspective and the opposite perspective. For these reasons, we took a different approach: we did not ask participants to identify their depression status, and we instructed all of them to complete a task from the perspective of a depressed and a non-depressed person. One aspect of their behaviour that we cannot measure using this approach is the knowledge of how participants would have behaved with their own perspective. We do, however, gain insight into their perspective-taking abilities, operationalized as a difference between two perspectives that they have to take. This insight is the focus of our study; we leave the exploration of the first approach to future research.

In what follows we formulate our hypotheses and outline an experimental task that enables these measurements.

First, if individuals can take on the perspective of someone with depression, they are expected to show affective behaviour similar to that found in depression. What does this behaviour look like? People who have depression experience a pessimism bias and predict negative outcomes for their future (Pyszczynski, Holt, & Greenberg, 1987), whereas people without depression are more optimistic (Sharot, 2011). But the exact characteristics of affective behaviour in depression are unclear. One of the most common symptoms is anhedonia, or the loss of pleasure: it leads to reduced positive affect and weaker responses to positive stimuli, known as positive attenuation (Epstein et al., 2006; Watson, Clark, & Carey, 1988). Other literature reports that those with depression have blunted emotion across the entire range, emotion context insensitivity, and their perception of the world is neither positive nor negative, but neutral (Bylsma, Morris, & Rottenberg, 2008; Rottenberg & Gotlib, 2004). Behaviourally, this surfaces as lethargy, fatigue, and lack of influence from external emotions. In other words, a "depressed" perspective might arguably cause responders to give weaker responses to positive words only (replicating positive attenuation) or to the entire word range (replicating emotion context insensitivity). Conversely, an assumed "non-depressed" perspective is expected to lead to a broader, unconstrained range of emotional responses.

Second, in line with the literature on traits similar to perspective-taking, such as Theory of Mind, we expect the responder's severity of depressive symptoms to modulate their perspective-taking ability. An oft-cited reduction in the Theory of Mind performance among depressed individuals, as well as a relatively pessimistic outlook associated with hopelessness suggest that depressive symptoms may come with inability to put oneself into another's shoes (Abramson et al., 1989; Y. Inoue et al., 2006; Yumiko Inoue, Tonooka, Yamada, & Kanba, 2004; Pyszczynski et al., 1987; Wang et al., 2008; Wolkenstein et al., 2011). Thus, we predict that, while taking different perspectives, an average individual with depressive symptoms will change their emotional responses to words to a lesser degree than an average individual without these symptoms.

An experimental paradigm that enables us to pursue these hypotheses within-subjects is that of a slider task (Warriner et al., 2017). The task consists of a scale and a humanoid manikin. The manikin is placed in the center of a (typically vertical) scale on a computer screen and can either represent the participant or someone else. Additionally, a word appears at the top or the bottom of the screen. Participants are instructed to move the manikin as close to, or as far away from, the word as they would prefer to be: they can be instructed to convey their own judgment or that of another person's perspective. The distance the manikin is from the word represents the affective rating of that word by the participant (smaller distance is higher valence; see Figure 1). Warriner et al. have found that the affective ratings from this task, when the referent was the participant, strongly correlated with affective ratings from a Likert scale, while (Imbault, Shore, & Kuperman, 2017) have demonstrated that the slider task is a reliable measure of valence for words both within and across individuals. The benefit of the slider method, compared to the ratings method, is that we can easily change the referent for the affective ratings, asking people to rate the emotion of words for someone who is not themselves. In the present task, all participants were asked to respond on a slider scale to a selection of words representing the entire spectrum of positivity and negativity, taking on the perspective of either a depressed or a non-depressed person. The outcomes allow for an easy quantification of the influences that perspective-taking and the presence or absence of depressive symptoms have on affective behaviour.

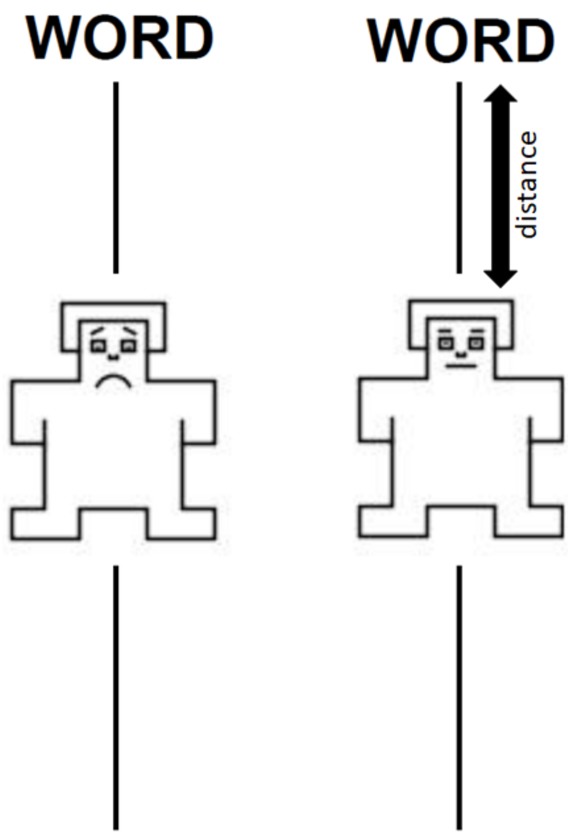

Figure 1 Illustration of the two conditions in the slider task with a sad (left) and neutral (right) manikins and how the distance to the word is calculated. Participants only see one manikin at any given time.

Results and Discussion

The initial raw data set consisted of 13,458 trials. Fifteen participants completed the task (at least partially) twice; we removed their second pass from analyses (2,808 trials), resulting in 10,650 trials. An additional 300 trials (2 participants) were removed from analysis due participants not completing the QIDS, resulting in 10,350 trials. We then, post-hoc, removed participants who fell outside of our extreme-groups QIDS criteria, outlined above. The resulting data pool contained 6,877 trials from 47 participants.

The Basic Empathy Scale can be broken down in two sub-scales; cognitive empathy and affective empathy. Overall empathy scores for the non-depressed group, which ranged from 59 to 95 (M = 75.923, SD = 10.103), were not significantly different from the depressed group, which ranged from 60 to 98 (M = 80.476, SD = 10.196) (t = -1.528, p = 0.134). Additionally, cognitive empathy scores for the non-depressed group, which ranged from 28 to 43 (M = 37.54, SD = 4.12), were not significantly different (t = -0.089, p = 0.93) from the depressed group, which ranged from 26 to 45 (M = 37.67, SD = 5.46). Interestingly, the depressed group had higher affective empathy scores than the non-depressed group (depressed: range 30:53, M = 42.81, SD = 7.35; non-depressed: range 26:52; M = 38.38, SD = 7.25; t = -2.06, $p < 0.05$). All empathy scores were not found to influence individual performance in the slider task, nor did they enter into an interaction with any critical variable: they are not reported further. We thank an anonymous reviewer for suggesting this analysis.

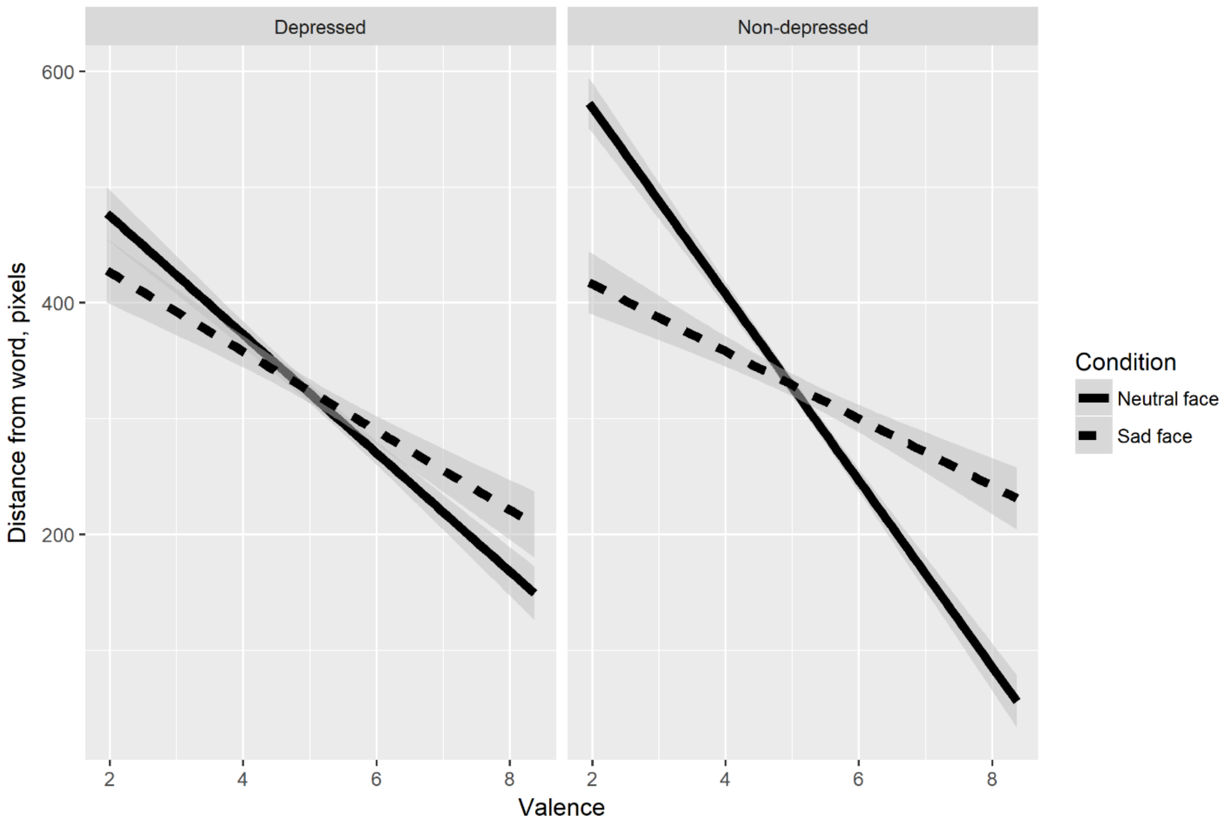

Figure 2. Distance from the word shown by depressed and non-depressed participants with both conditions (sad and neutral face)

A regression model with the three-way interaction of Face x Group x Valence quantifies changes in affective behaviour (manikin's distance to target word) caused by the assumed perspective, word valence and the severity of depressive symptoms, see Table 1. Figure 2 visualizes the response patterns by presenting the effect of valence on distance for a depressed group (left panel) and a non-depressed group (right panel) and for a sad and neutral manikin (dashed and solid lines respectively). The plot is based on linear fits to raw data: visual inspection of a plot with model predictions reveal that they are a close fit to Figure 2. In what follows, we describe patterns of interest based either on the model with a three-way interaction in Table 1 or the models that highlight specific contrasts by using two-way interactions nested in that bigger model (not shown).

In all conditions, greater word valence came with a shorter distance between the manikin and the word. In line with all prior reports (Imbault et al., 2017; Warriner et al., 2017), positive emotion is associated with an approach behaviour. However, marked differences were observed both across groups and across experimental conditions in how strongly the range of positivity influenced responses to words. The steepest slope (and the stronger effect of valence and range of responses) was found in the non-depressed group taking a neutral perspective (*B* = -79.58, which is a measure of change in pixels per 1 unit of valence). The respective slope values for depressed participants with neutral manikin, non-depressed participants with sad manikin and depressed participants with sad manikin were *B* = -51.44, -26.23 and -33.46.

These data enable us to answer the questions posited in the Introduction. Our first analysis compares the change in perspective, arising from completing the task with a sad face versus a neutral face (i.e., mimicking a non-depressed versus a depressed perspective). Both groups changed their behaviour in a way that showed a blunter affect with a sad, depressed perspective (dashed lines had flatter slopes than solid ones in both panels). Yet in the non-depressed group the change in perspective had a much stronger effect on distance than in the depressed group (non-depressed: B = 52.987, SE = 4.637, t = 11.428, $p < 0.001$; depressed: B = 18.309, SE = 5.153, t = -3.553, $p<0.001$). This cross-group difference between conditions was significant, as the three-way Face x Group x Valence interaction showed (B = -35.487, SE = 6.975, t = -5.088,

$p < 0.001$). Taken together, these findings indicate that all participants are capable of perspective-taking, yet this ability is weakened in people with depressive symptoms.

Predictor	Estimate	*SE*	*t*-value	*p* value
Intercept	322.309	8.753	36.823	<0.001
Valence	-33.803	12.319	-2.744	0.008
Non-depressed Group	2.730	11.185	0.244	0.808
“Neutral” Face	-9.208	6.345	-1.451	0.147
Valence: Non-depressed Group	5.508	16.297	0.338	0.737
Valence: “Neutral” Face	-18.036	5.104	-3.534	<0.001
Non-depressed Group: “Neutral” Face	0.242	8.637	0.028	0.977
Valence: Non-depressed Group: “Neutral” Face	-35.487	6.975	-5.088	<0.001

Table 1: Linear mixed-effects model fitted to the distance of the manikin from the word. N = 6858. SD of random by-word intercepts = 33.55, SD of random by-subject intercepts = 32.08, SD of by-subject random slopes of valence = -0.21, SD of residual = 173.30. Reference levels are “depressed” for Group, and “Sad” for Face.

How good were participants in changing their perspective? Individuals in the non-depressed group were extremely accurate in approximating the behaviour of individuals with depressive symptoms when both groups were moving the sad manikin (dashed lines across panels in Figure 2). There was no reliable contrast in the valence slopes for these two groups when operating the sad manikin ($B = 5.686$, $SE = 16.369$, $t = 0.347$, $p = 0.730$). However, the opposite was not true. Individuals with depressive symptoms were unable to change their perspective and replicate the wide range of emotional responses characteristic of non-depressed individuals when both groups moved the neutral manikin. The difference between the two groups' slopes in the neutral-face condition was marginally significant ($B = 29.889$, $SE = 16.241$, $t = 1.840$, $p = 0.07$), with the depressed group's slope being 29.89 pixels (or 4.5% of the available distance scale) flatter than the non-depressed group per unit of valence (solid lines across panels in Figure 2). Across the entire range of valence, this difference equates to 29.57% of the available scale.

Finally, the patterns demonstrate that either depressive symptoms or a depressed perspective have impact on the entire range of valence, and not specifically positive or negative stimuli. A hypothetical regression line that would show a perfect symmetry between impacts on positive and negative words is predicted to have an intercept of 325 pixels (the middle of the slider scale) at the mean value of the valence range. In fact, all regression lines were nearly perfectly centered: all intercepts were within 10 pixels of the mid-point of the slider scale, see Figure 2 and relevant contrasts in Table 1. This finding sheds light on a long-standing debate of whether the pessimistic outlook only influences the positive side of the emotional spectrum, its negative side, or the entire spectrum: the latter is true.

General Discussion

This study pursued two novel questions regarding emotional states of depressed and non-depressed individuals and their respective abilities to take each other's perspective. First, we quantified between participants how the presence or absence of depressive symptoms influenced the magnitude of their responses to words representing an entire range from very negative to very positive. Second, we measured within-participants how their responses change as a function of the perspective they are assigned to take, that of a depressed or a non-depressed person. Our main interest is in the interaction of the two effects: we examine how one's emotional state determines the ability to imagine someone else's responses.

Unsurprisingly, individuals with depressive symptoms respond differently to emotional stimuli than those without depression. In line with previous literature (Bylsma et al., 2008; Epstein et al., 2006; Rottenberg & Gotlib, 2004; Watson et al., 1988), when the participants with depressive symptoms operated the sad manikin, they did not use the entire range of the slider scale, but instead remained relatively neutral to all levels of valence. One of our questions concerned whether blunted emotion was specific to positive attenuation (Epstein et al., 2006; Watson et al., 1988) and thus restricted to anhedonia, or the entire scale, expressed as emotion context insensitivity (Bylsma et al., 2008; Rottenberg & Gotlib, 2004). Our data offer a clear demonstration of the latter hypothesis, in support of a notion that an across-the-board flattening of affect is symptomatic of depression (American Psychiatric Association, 2013). On the other hand, and consistent with previous research with the slider paradigm, participants without depression were strongly affected by valence and used the entire range of the slider scale in response to valence when moving a neutral manikin (Imbault et al., 2017; Warriner et al., 2017).

A more intriguing new finding is the effect of depression on perspective-taking abilities. The findings that those with depression have heightened empathy and lowered Theory of Mind are well established in the literature (Farrow & Woodruff, 2008; Haaga & Beck, 1995; Yumiko Inoue et al., 2004; O'Connor et al., 2007; Wang et al., 2008; Wolkenstein et al., 2011), but there has been little to no research on how depression would impact perspective taking abilities. By manipulating the slider paradigm, i.e., by asking those with depression to behave like people without depression, we were able to quantify this impact. The comparison was achieved by the analysis of the slider task performance across depressed and non-depressed groups when operating a neutral manikin. When answering for another group, the participants with depressive symptoms did not match the large effect that valence had on distance for that (non-depressed) group. Instead they responded with a more shallow slope, i.e., distance from the word was less influenced by valence. These responses in the depressed group were similar to those they produced when operating a sad manikin. We take this to mean that our participants with depression had trouble with taking the perspective of a person without depression. An alternative

interpretation of these results is that participants with depressive symptoms have overall smaller emotional reactivity, and their observed inability to take on other perspective is just an inability to expand their emotional range. Although it cannot be ruled out given the data, the experimental design does not warrant this interpretation. Perspective-taking is purely a mental exercise; we do not ask participants to *feel* these emotions, rather we ask them to speculate how another group feels these emotions. Thus, their own reduced emotional reactivity per se is not expected to prevent participants with depressive symptoms from hypothesizing what emotional responses would be plausible under a different perspective. Thus, arguably, an inability to take on another perspective is at least partly independent from one's range of emotional reactivity.
Our findings are consistent with such theoretical accounts as the Hopelessness Theory of Depression (Abramson et al., 1989) and the pessimism bias (Pyszczynski et al., 1987). An inability to understand what it feels like to not be depressed is a harrowing finding, and leads into a "no light at the end of the tunnel" feeling of hopelessness. Autonomous motivation, i.e. "I want to get better", is a strong predictor of the outcome of depression, resulting in higher remission and lower depression severity (Zuroff et al., 2007). Our findings indicate that a critical component for this motivation is missing in the outlook of a person with depressive symptoms: namely, an understanding of what being better feels like.

Another new finding is that non-depressed participants were able to accurately adopt the perspective of the feeling of depression. When these participants moved the sad manikin, they accurately approximated the much weaker effect of valence on the distance from the words. The observation that emotional states linked to depressive symptoms are accessible to the majority of the population is valuable for the prognosis of depression. As shown in studies of this mood disorder, having a strong social network and feeling understood leads to faster recovery (Cutrona & Troutman, 1986; Eisman et al., 2015; George et al., 1989; Peirce et al., 2000).

To our knowledge, this is the first time when perspective-taking ability has been directly tested within subjects and in the context of depressive symptomatology. We observed in two groups with and without severe depressive symptoms highly distinct patterns of affective behaviour, which were contingent on both their own symptomatology and who they represented in the task. Depressive symptoms restricted an emotional range to the point where a "normal" perspective appeared unattainable. Conversely, depressive mood was an easy target for people who were not in that emotional state. This set of findings is noteworthy because of its potential implications for diagnostics and therapy of depression. It is also notable because it was obtained with a very simple online experimental tool in a non-invasive way and yet gave rise to strong and statistically reliable differences between groups and within individuals. The new slider method that we introduced offers easy quantification of every aspect of participants' behavior and the independent and interactive contributions of all manipulated factors. As such, this tool enables easy replication in similar non-clinical cohorts and, importantly, a much-needed further expansion over clinical populations. The slider task can also quantify perspective-taking abilities in many groups, including those varying in gender, age, cultural background, or indeed their clinical status.

Methods

Participants

We recruited participants from a pool independently created by the BrainTrain project (Waris et al., 2017) using the crowd-sourcing platform Amazon Mechanical Turk (www.mturk.com). This study was approved by the McMaster Research Ethics Board, protocol 2011 165, informed consent was obtained from all participants explaining any risks involved, and all methods were carried out in accordance with relevant guidelines and regulations. 71 participants from the pool agreed to take part in our study and were compensated monetarily. All but 2 participants had completed QIDS-SR prior to being recruited for this experiment (The Quick Inventory of Depressive Symptomatology; a self-report, 16-item questionnaire about the symptoms of depression; 27) ; the data for two participants were removed from analysis. We adopted the clinical cut-offs of the QIDS-SR scale, and defined a non-depressed group as one with QIDS-SR score < 6 and a group with moderate to severe symptoms of depression as one with QIDS-SE score ≥ 11. We removed 22 participants that did not meet the cut-offs of our extreme groups design.

The remaining 47 participants (23 female) ranged in age from 19 to 72 (*M* = 34.74, *SD* = 13.24). The 26 participants (9 female) in the "non-depressed" group ranged in age from 19 to 72 (*M* = 37.27, *SD* = 15.61). Their QIDS-SR scores ranged from 0 to 5 (*M* = 2.385, *SD* = 1.65). The 21 participants (14 female) in the "depressed" group ranged in age from 20 to 50 (*M* = 31.62, *SD* = 8.94). Their QIDS-SR scores ranged from 11 to 22 (*M* = 14.10, *SD* = 2.96).

It is important to note that our participants were selected based on their QIDS-SR score, which is a self-reported questionnaire about the *symptoms* of depression. We have no information on whether our participants actually had clinical depression. Thus, our group labels "depressed" and "non-depressed" refer to the *symptoms* exhibited at the time of self-report, and not necessarily to the presence of the mood disorder. Participants were not asked whether they self-identified with having a depression (see motivation above). However, we made a simplifying assumption that the affective behaviour shown by participants with depressive symptoms when taking a "sad, depressed" perspective is representative of their own behaviour, much like the behaviour shown by non-depressed individuals with a neutral manikin is indicative of their own behaviour.

Affective Stimuli

A set of 150 words were selected to represent an entire range of psychological valence (positivity and negativity) and arousal. The norms of valence and arousal are available in Warriner et al.'s (Warriner et al., 2013) and are evaluated on a scale from 1 (sad or calm) to 9 (happy or excited). First, we chose monosyllabic words from Warriner et al.'s(Warriner et al., 2013) list of 13,763 words. These words were divided into 25 bins (5 quintiles of valence x 5 quintiles of arousal), and 6 words were randomly selected from each bin, guaranteeing a relatively uniform coverage of the affective range. The words had an average length of 4.34 characters (*SD* = 0.92). The mean frequency of occurrence in the 51 million-token corpus of subtitles to the US films and media SUBTLEX was 2780.53 (*SD* = 6338.43) (Brysbaert & New, 2009). Frequency was not correlated with arousal (r_s = 0.05, p = 0.51), but was correlated with valence (r_s = 0.31, p < 0.01). Valence and arousal were not correlated (r_s = -0.05, p = 0.54).

Procedure

Participants completed the slider task on a web application that was implemented using jQuery and PHP. They used their own computers and were required to do so in Full Screen mode. Participants could use either a trackpad or a mouse to perform the experiment.

Prior to the start of the experiment, participants filled out demographic information (including age, sex, handedness, current US state and native languages). Immediately after, participants completed the Basic Empathy Scale (Jolliffe & Farrington, 2006) which includes 20 questions on a 5-point Likert scale ("Strongly Disagree" to "Strongly Agree") that measures both cognitive and affective empathy.

After completing the Basic Empathy Scale, participants read the instructions for the slider task: *"[…] you will see a word either at the top or the bottom with a vertical line below or above it. There will be a person in the centre of that line. The person represents either a neutral person or a sad, depressed person. You can move the person closer to or further away from the word. Position the person where you think they would prefer to be."*

The "person" (manikin) was positioned in centre of the line, with a word at the top or the bottom of the screen (see Figure 1). This task is an adaption of Warriner et al.(Warriner et al., 2017). Rather than asking participants to respond to the emotion of word for themselves, we asked the participants to respond as either a "sad, depressed" person, or a "neutral" person.

After reading the instructions, participants completed 5 practice trials with both types of manikin before continuing the experiment. The affective manikins were selected from Bradley & Lang (Bradley & Lang, 1999); the neutral manikin is the middle manikin from the valence scale, and the sad manikin is the lowest valence manikin from the valence scale. The emotion of the manikin was indicated at the beginning of each block and was confined to the shape of the mouth and the eyebrows (straight line for both in the neutral manikin, and raised eyebrows and a downturned mouth in the sad one, see Figure 1). The manikin started in the centre position of the scale; 325 pixels away from the word. The participants were able to move the manikin as many times as they felt necessary before clicking the Submit button, but only the final position of the manikin was used in our data analysis. Every participant saw all 150 words, and word order was randomized within each block. Order of the blocks were counter balanced; half of the participants saw the sad manikin first, while the other half saw the neutral manikin first. All participants saw the same 75 words for the sad manikin and the same 75 words for the neutral manikin.

Variables

The dependent variable of interest was the distance of the manikin from the word (measured in pixels) when the Submit button was pressed (labeled as Distance). The distance occupied a range of 648 pixels, 1 being closest to the word and 649 being furthest from the word.

One independent variable of interest is the word's valence rating (Valence), which was retrieved from Warriner, Kuperman, & Brysbaert(Warriner et al., 2013). The valence ratings were the mean response of 20 raters given on a 1 (unhappy) to 9 (very happy) scale. An additional variable of interest was how empathetic the participants was, based on their Basic Empathy Scale score (Jolliffe & Farrington, 2006). Another critical variable was Group, a between-subjects binary indicator of a group label (depressed or non-depressed) to which a participant belonged, based on the QIDS score (Rush et al., 2003). Finally, a within-subjects binary indicator of the manikin type Face (sad vs neutral) marked the perspective that the participant was instructed to take. We expected the effects of interest to emerge in three or four-way interactions of these critical variables.

Statistical Analyses

We used linear mixed-effects multiple regression models with participants and words as crossed random effects (Baayen, Davidson, & Bates, 2008; Soares et al., 2012), as implemented in package lmerTest version 2.0-33 (Kuznetsova, Brockhoff, & Christensen, 2016) for R version 3.4.1 (R Core Team, 2017). This method enables a simultaneous exploration of multiple factors and covariates, while accounting for between-participants and between-items variance. Each model was initially fitted with random intercepts for words and participants and by-participant random slopes of valence, and trimmed down to only contain the random effects that significantly improve the model's performance, as indicated by a series of likelihood ratio tests that compared a model with a given random effect and a model without this random effect. Using the same test in the backward elimination procedure, we removed all fixed effects that did not improve the model's performance. To reduce the influence of outliers, the frequency estimates were (natural) log-transformed, as indicated by the Box-Cox power transformation test. After each fit, we excluded any data with residuals that fell 2.5 standard deviations away from the mean and refitted the trimmed models. Results are visualized with the help of the ggplot2 package (Wickham & Chang, 2016). Contrasts were coded with dummy coding.

Data Availability

The datasets analysed during the current study are available from the corresponding author (imbaulcl@mcmaster.ca) on request.

References

Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review*, *96*(2), 358–372. https://doi.org/10.1037/0033-295X.96.2.358

Albaum, G., Best, R., & Hawkins, D. E. L. (1981). Continuous vs Discrete Semantic Differential Rating Scales. *Psychological Reports*, *49*, 83–86.

Allen, N. B., Trinder, J., & Brennan, C. (1999). Affective Startle Modulation in Clinical Depression: Preliminary Findings. *Biological Psychiatry*, *46*(4), 542–550. https://doi.org/https://doi.org/10.1016/S0006-3223(99)00025-6

Altarriba, J., & Bauer, L. M. (2004). The Distinctiveness of Emotion Concepts: A Comparison between Emotion, Abstract, and Concrete Words. *The American Journal of Psychology*, *117*(3), 389–410.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (DSM-5®). *American Psychiatric Pub*.

Anooshian, L. J., & Hertel, P. T. (1994). Emotionality in free recall: Language specificity in bilingual memory. *Cognition and Emotion*, *8*(6), 503–514. https://doi.org/10.1080/02699939408408956

Ayçiçeği-Dinn, A., & Caldwell-Harris, C. L. (2009). Emotion-memory effects in bilingual speakers: A levels-of-processing approach*. *Bilingualism: Language and Cognition*, *12*(03), 291. https://doi.org/10.1017/S1366728909990125

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(04), 577–660. https://doi.org/10.1017/S0140525X99002149

Batson, C. D., Early, S., & Salvarani, G. (1997). Perspective Taking: Imagining How Another Feels Versus Imaging How You Would Feel. *Personality and Social Psychology Bulletin*.

https://doi.org/10.1177/0146167297237008

Bayer, M., Sommer, W., & Schacht, A. (2011). Emotional words impact the mind but not the body: Evidence from pupillary responses. *Psychophysiology*, *48*(11), 1554–1562. https://doi.org/10.1111/j.1469-8986.2011.01219.x

Betella, A., & Verschure, P. F. M. J. (2016). The affective slider: A digital self-assessment scale for the measurement of human emotions. *PLoS ONE*, *11*(2), 1–11. https://doi.org/10.1371/journal.pone.0148037

Bonin, P., Aubert, L., Malardier, N., & Niedenthal, P. M. (2016). Subjective Et De Valence Émotionnelle Pour 866 Mots, *103*, 655–694.

Bradley, M. M., & Lang, P. P. J. (1999). Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings. *Psychology*, *Technical*(C–1), 0. https://doi.org/10.1109/MIC.2008.114

Brase, J., & Mani, N. (2017). Effects of learning context on the acquisition and processing of emotional words in bilinguals. *Emotion*, *17*(4), 628–639. https://doi.org/10.1037/emo0000263

Briesemeister, B. B., Kuchinke, L., & Jacobs, A. M. (2011). Discrete emotion effects on lexical decision response times. *PLoS ONE*, *6*(8). https://doi.org/10.1371/journal.pone.0023743

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990. https://doi.org/10.3758/BRM.41.4.977

Bylsma, L. M., Morris, B. H., & Rottenberg, J. (2008). A meta-analysis of emotional reactivity in major depressive disorder. *Clinical Psychology Review*, *28*(4), 676–691. https://doi.org/10.1016/j.cpr.2007.10.001

Caldwell-Harris, C. L. (2015). Emotionality Differences Between a Native and Foreign Language. *Current Directions in Psychological Science*, *24*(3), 214–219. https://doi.org/10.1177/0963721414566268

Caldwell-Harris, C. L., Tong, J., Lung, W., & Poo, S. (2011). Physiological reactivity to emotional phrases in Mandarin--English bilinguals. *International Journal of Bilingualism*, *15*(3), 329–352. https://doi.org/10.1177/1367006910379262

Canli, T., Sivers, H., Whitfield, S. L., Gotlib, I. H., & Gabrieli, J. D. E. (2002). Amygdala response to happy faces as a function of extraversion. *Science*, *296*(5576), 2191. https://doi.org/10.1126/science.1068749

Cartwright, C., Gibson, K., Read, J., Cowan, O., & Dehar, T. (2016). Long-term antidepressant use: patient perspectives of benefits and adverse effects. *Patient Preference and Adherence*, *Volume 10*, 1401–1407. https://doi.org/10.2147/PPA.S110632

Carver, C. S., & White, T. L. (1994). Behavioral Inhibition, Behavioral Activation and affective responses to impending reward and punishment. *Journal of Personality and Social Psychology*, *67*, 319–333.

Chen, M., & Bargh, J. A. (1997). Nonconscious behavioral confirmation processes: The self-fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology*, *33*(5), 541–560. https://doi.org/10.1006/jesp.1997.1329

Citron, F. M. M., Weekes, B. S., & Ferstl, E. C. (2013). Effects of valence and arousal on written word recognition: Time course and ERP correlates. *Neuroscience Letters*, *533*, 90–95.

Clark, J., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*(3), 149–210. https://doi.org/10.1007/BF01320076

Cutrona, C. E., & Troutman, B. R. (1986). Social Support , Infant Temperament , and Parenting Self-Efficacy : A Mediational Model of Postpartum Depression Author (s): Carolyn E . Cutrona and Beth R . Troutman Published by : Wiley on behalf of the Society for Research in Child Development Stable, *57*(6), 1507–1518.

Daneman, M., & Carpenter, P. A. (1980). Individual Differences in Working Memory and Reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466.

De Grauwe, S., Willems, R. M., Rueschemeyer, S. A., Lemhöfer, K., & Schriefers, H. (2014). Embodied language in first- and second-language speakers: Neural correlates of processing motor verbs. *Neuropsychologia*, *56*(1), 334–349.

https://doi.org/10.1016/j.neuropsychologia.2014.02.003

De Pascalis, V., Fiore, A. D., & Sparita, A. (1996). Personality, event-related potential (ERP) and heart rate (HR): An investigation of Gray's theory. *Personality and Individual Differences*, *20*(6), 733–746. https://doi.org/10.1016/0191-8869(96)00016-5

De Pascalis, V., Strippoli, E., Riccardi, P., & Vergari, F. (2004). Personality, event-related potential (ERP) and heart rate (HR) in emotional word processing. *Personality and Individual Differences*, *36*(4), 873–891. https://doi.org/10.1016/S0191-8869(03)00159-4

Decety, J., & Batson, C. D. (2007). Social neuroscience approaches to interpersonal sensitivity. *Social Neuroscience*, *2*(3–4), 151–157. https://doi.org/10.1080/17470910701506060

Degner, J., Doycheva, C., & Wentura, D. (2012). It matters how much you talk: On the automaticity of affective connotations of first and second language words. *Bilingualism*, *15*(1), 181–189. https://doi.org/10.1017/S1366728911000095

Dewaele, J.-M. (2004). The Emotional Force of Swearwords and Taboo Words in the Speech of Multilinguals. *Journal of Multilingual and Multicultural Development*, *25*(23), 204–222. https://doi.org/10.1080/01434630408666529

Dewaele, J. M., & Nakano, S. (2013). Multilinguals' perceptions of feeling different when switching languages. *Journal of Multilingual and Multicultural Development*, *34*(2), 107–120. https://doi.org/10.1080/01434632.2012.712133

Dresler, T., Mériau, K., Heekeren, H. R., & Van Der Meer, E. (2009). Emotional Stroop task: Effect of word arousal and subject anxiety on emotional interference. *Psychological Research*, *73*(3), 364–371. https://doi.org/10.1007/s00426-008-0154-6

Ehrman, M. E., Leaver, B. Lou, & Oxford, R. L. (2003). A brief overview of individual differences in second language learning. *System*, *31*(3), 313–330. https://doi.org/10.1016/S0346-251X(03)00045-9

Eilola, T. M., Havelka, J., & Sharma, D. (2007). Emotional activation in the first and second language. *Cognition and Emotion*, *21*(5), 1064–1076. https://doi.org/10.1080/02699930601054109

Eisman, A. B., Stoddard, S. a, Heinze, J., Caldwell, C. H., & Zimmerman, M. a. (2015). Depressive symptoms, social support, and violence exposure among urban youth: A longitudinal study of resilience. *Developmental Psychology*, *51*(9), 1307–1316. https://doi.org/10.1037/a0039501

Ekman, P. (1992). An Argument for Basic Emotions. *Cognition and Emotion, 6*(3–4), 169–200. https://doi.org/10.1080/02699939208411068

Elliot, A. J. (2006). The Hierarchical Model of Approach-Avoidance Motivation. *Motivation and Emotion, 30*(2), 111–116. https://doi.org/10.1007/s11031-006-9028-7

Epstein, J., Pan, H., Kocsis, J., Yang, Y., Butler, T., Chusid, J., … Silbersweig, D. (2006). Lack of Ventral Striatal Response to Positive Stimuli inDepressed Versus Normal Subjects. *The American Journal of Psychiatry*, 1784. https://doi.org/10.1176/ajp.2006.163.10.1784

Farrow, T. F. D., & Woodruff, P. W. R. (2008). Empathy in Mental Illness. *Psychiatric Services*, *59*, 581–581. https://doi.org/10.1176/appi.ps.59.5.581

Feldman Barrett, L., & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology*, *74*(4), 967–984. https://doi.org/10.1037/0022-3514.74.4.967

Ferré, P., García, T., Fraga, I., Sánchez-Casas, R., & Molero, M. (2010). Memory for emotional words in bilinguals: Do words have the same emotional intensity in the first and in the second language? *Cognition and Emotion*, *24*(5), 760–785. https://doi.org/10.1080/02699930902985779

Foroni, F. (2015). Do we embody second language? Evidence for ‘partial’ simulation during processing of a second language. *Brain and Cognition*, *99*, 8–16. https://doi.org/10.1016/j.bandc.2015.06.006

George, L., Blazer, D., Hughes, D., & Fowler, N. (1989). Social support and the outcome of major depression. *British Journal of Psychiatry*, *154*, 478–485.

Golin, S., Hartman, S. A., Klatt, E. N., Munz, K., & Wolfgang, G. L. (1977). Effects of Self-esteem Manipulation on Arousal and Reactions to Sad Models in Depressed and Nondepressed College Students. *Journal of Abnormal Psychology*, *86*(4), 435–439.

Haaga, D. A. F., & Beck, A. T. (1995). Perspectives on depressive realism: Implications for cognitive theory of depression. *Behaviour Research and Therapy*, *33*(1), 41–48. https://doi.org/10.1016/0005-7967(94)E0016-C

Harris, C. L., Ayçiçeği, A., & Gleason, J. B. (2003). Taboo words and reprimands elicit greater autonomic reactivity in a first language than in a second language. *Applied Psycholinguistics*, *24*(4), 561–579. https://doi.org/10.1017/S0142716403000286

Harris, C. L., Gleason, J. B., & Ayçiçeği, A. (2006). When is a First Language More Emotional? Psychophysiological Evidence from Bilingual Speakers. In *Bilingual minds: Emotional experience, expression and representation* (pp. 257–283). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.69.5830%0Ahttp://www.zeit.de/wirtschaft/2014-08/russland-import-sanktionen%0Ahttp://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.69.5830%0Ahttp://www.zeit.de/wirtschaft/2014-08/russland-import-sa

Hellman, A. B. (2011). Vocabulary size and depth of word knowledge in adult-onset second language acquisition. *International Journal of Applied Linguistics*, *21*(2), 162–182. https://doi.org/10.1111/j.1473-4192.2010.00265.x

Hinojosa, J. a., Martínez-García, N., Villalba-García, C., Fernández-Folgueiras, U., Sánchez-Carmona, a., Pozo, M. a., & Montoro, P. R. (2015). Affective norms of 875 Spanish words for five discrete emotional categories and two emotional dimensions. *Behavior Research Methods*, (August). https://doi.org/10.3758/s13428-015-0572-5

Imbault, C., Shore, D., & Kuperman, V. (2017). Reliability of the sliding scale for collecting affective responses to words. *Behavior Research Methods*.

Inoue, Y., Yamada, K., & Kanba, S. (2006). Deficit in theory of mind is a risk for relapse of major depression. *Journal of Affective Disorders*, *95*(1–3), 125–127. https://doi.org/10.1016/j.jad.2006.04.018

Inoue, Yumiko, Tonooka, Y., Yamada, K., & Kanba, S. (2004). Deficiency of theory of mind in patients with remitted mood disorder. *Journal of Affective Disorders*, *82*(3), 403–409. https://doi.org/10.1016/j.jad.2004.04.004

Itkes, O., Eviatar, Z., & Kron, A. (2019). Semantic and affective manifestations of ambi (valence). *Cognition and Emotion*, *9931*. https://doi.org/10.1080/02699931.2018.1564249

Itkes, O., Kimchi, R., Kron, A., & Carmel, M. (2017). Dissociating Affective and Semantic Valence, *146*(7), 1–36. https://doi.org/10.1037/xge0000291

Itzhak, I., Vingron, N., Baum, S. R., & Titone, D. (2017). Bilingualism in the real world: How proficiency, emotion, and personality in a second language impact communication in clinical and legal settings. *Translational Issues in Psychological Science*, *3*(1), 48–65. https://doi.org/10.1037/tps0000103

Johnson-Laird, P. N., & Oatley, K. (1989). The Language of Emotions: An Analysis of a Semantic Field. *Cognition and Emotion*, *3*(2), 81–123. https://doi.org/10.1080/02699938908408075

Jolliffe, D., & Farrington, D. P. (2006). Development and validation of the Basic Empathy Scale. *Journal of Adolescence*, *29*(4), 589–611. https://doi.org/10.1016/j.adolescence.2005.08.010

Just, M. A., & Carpenter, P. A. (1999). A Capacity Theory of Comprehension : Indicidual Dirrefences in Working Memory, (January), 1–16.

Kazanas, S. A., & Altarriba, J. (2016). Emotion Word Processing: Effects of Word Type and Valence in Spanish–English Bilinguals. *Journal of Psycholinguistic Research*, *45*(2), 395–406. https://doi.org/10.1007/s10936-015-9357-3

Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual Differences in Language Acquisition and Processing. *Trends in Cognitive Sciences*, *22*(2), 154–169. https://doi.org/10.1016/j.tics.2017.11.006

Kleinginna, P. R., & Kleinginna, A. M. (1981). A categorized list of motivation definitions, with a suggestion for a consensual definition. *Motivation and Emotion*, *5*(3), 263–291. https://doi.org/10.1007/BF00993889

Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The Representation of Abstract Words: Why Emotion Matters. *Journal of Experimental Psychology: General*, *140*(1), 14–34. https://doi.org/10.1037/a0021446

Kousta, S. T., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition*, *112*(3), 473–481. https://doi.org/10.1016/j.cognition.2009.06.007

Kuchinke, L., Võ, M. L. H., Hofmann, M., & Jacobs, A. M. (2007). Pupillary responses during lexical decisions vary with word frequency but not emotional valence. *International Journal of Psychophysiology*, *65*(2), 132–140. https://doi.org/10.1016/j.ijpsycho.2007.04.004

Kunzmann, U., & Grühn, D. (2005). Age differences in emotional reactivity: The sample case of sadness. *Psychology and Aging*, *20*(1), 47–59. https://doi.org/10.1037/0882-7974.20.1.47

Kuperman, V., Estes, Z., Brysbaert, M., & Warriner, A. B. (2014). Emotion and language: Valence and arousal affect word recognition. *Journal of Experimental …*, (4). Retrieved from http://psycnet.apa.org/psycarticles/2014-03909-001

Kuperman, V., & Van Dyke, J. a. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language*, *65*(1), 42–73. https://doi.org/10.1016/j.jml.2011.03.002

Kuppens, P., Tuerlinckx, F., Yik, M., Koval, P., Coosemans, J., Zeng, K. J., & Russell, J. A. (2017). The Relation Between Valence and Arousal in Subjective Experience Varies With Personality and Culture. *Journal of Personality*, *85*(4), 530–542. https://doi.org/10.1111/jopy.12258

Kuznetsova, A., Brockhoff, P., & Christensen, R. (2016). lmerTest: Tests in Linear Mixed Effects Models. *R Package Version*, *3.0.0,* https://cran.r-project.org/package=lmerTest. Retrieved from https://cran.r-project.org/package=lmerTest

Lang, P. J. (1995). The Emotion Probe: Studies of Motivation and Attention. *American Psychologist*. https://doi.org/10.1037/0003-066X.50.5.372

Larsen, R. J., & Ketelaar, T. (1991). Personality and susceptibility to positive and negative emotional states. *Journal of Personality and Social Psychology*, *61*(1), 132–140. https://doi.org/10.1037/0022-3514.61.1.132

Levenson, R. W., Carstensen, L. L., Friesen, W. V., & Ekman, P. (1991). Emotion, physiology, and expression in old age. *Psychology and Aging*, *6*(1), 28–35.

https://doi.org/10.1037/0882-7974.6.1.28

Lewinsohn, P. M., Lobitz, W. C., & Wilson, S. (1973). "Sensitivity" of Depressed Individuals to Aversive Stimuli. *Journal of Abnormal Psychology*, *81*(3), 259–263.

Long, L. S., & Titone, D. A. (2007). Psychopathy and verbal emotion processing in non-incarcerated males. *Cognition and Emotion*, *21*(1), 119–145. https://doi.org/10.1080/02699930600551766

Mackay, D. G., Shafto, M., Taylor, J. K., Marian, D. E., Abrams, L., & Dyer, J. R. (2004). Relations between emotion, memory, and attention: Evidence from taboo Stroop, lexical decision, and immediate memory tasks. *Memory & Cognition*, *32*(3), 474–488. https://doi.org/10.3758/BF03195840

Macleod, J. W., Lawrence, M. a, McConnell, M. M., Eskes, G. a, Klein, R. M., & Shore, D. I. (2010). Appraising the ANT: Psychometric and theoretical considerations of the Attention Network Test. *Neuropsychology*, *24*(5), 637–651. https://doi.org/10.1037/a0019803

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research : JSLHR*, *50*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

Mcwhirter, B. T., Besett-Alesch, T. M., Horibata, J., & Gat, I. (2002). Loneliness in High Risk Adolescents: The Role of Coping, Self-Esteem, and Empathy. *Journal of Youth Studies*, *5*(1), 69–84. https://doi.org/10.1080/1367626012011177

Méndez-Bértolo, C., Pozo, M. A., & Hinojosa, J. A. (2011). Word frequency modulates the processing of emotional words: Convergent behavioral and electrophysiological data. *Neuroscience Letters*, *494*(3), 250–254. https://doi.org/10.1016/j.neulet.2011.03.026

Monnier, C., & Syssau, A. (2014). Affective Norms for French Words (FAN). *Behavior Research Methods*, *46*(4), 1128–1137. https://doi.org/10.3758/s13428-013-0431-1

Moors, A., De Houwer, J., Hermans, D., Wanmaker, S., van Schie, K., Van Harmelen, A.-L., … Brysbaert, M. (2013). Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words. *Behavior Research Methods*, *45*(1), 169–177.

https://doi.org/10.3758/s13428-012-0243-8

Mor, N., & Winquist, J. (2002). Self-focused attention and negative affect: A meta-analysis. *Psychological Bulletin*, *128*(4), 638–662. https://doi.org/10.1037/0033-2909.128.4.638

O'Connor, L. E., Berry, J. W., Lewis, T., Mulherin, K., & Crisotomo, P. S. (2007). Empathy and depression: the moral system on overdrive. *Empathy in Mental Ilness*. https://doi.org/10.1017/CBO9780511543753.005

Orne, M. T. (1962). On the Social Psychology of Psychological Experiment: With Particular Reference to Demand Characteristics and Their Implications. *American Psychologist*, *17*(11), 776–783. Retrieved from https://journals.scholarsportal.info/details/0003066x/v17i0011/776_otspottdcati.xml

Ostaszewski, K., & Zimmerman, M. A. (2006). The effects of cumulative risks and promotive factors on urban adolescent alcohol and other drug use: A longitudinal study of resiliency. *American Journal of Community Psychology*, *38*(3–4), 237–249. https://doi.org/10.1007/s10464-006-9076-x

Pearson, C., Janz, T., & Ali, J. (2013). *Mental and substance use disorders in Canada*.

Peirce, R., Frone, M., Russell, M., Cooper, M., & Mudar, P. (2000). A Longitudinal Model of Social Contact, Social Support, Depression, and Alcohol Use. *Health Psychol*, *19*(1), 28–38. https://doi.org/10.1037/0278-6133.19.1.28

Pennebaker, J., Booth, R., & Francis, M. (2007). Linguistic inquiry and word count: LIWC. Austin, TX: liwc net.

Perfetti, C. (2007). Reading Ability: Lexical Quality to Comprehension. *Scientific Studies of Reading*, *11*(4), 357–383. https://doi.org/10.1080/10888430701530730

Ponari, M., Rodriguez-Cuadrado, S., & Vinson, D. (2014). Processing advantage for emotional words in bilingual speakers. *Emotion*, *15*(January), 644–652. https://doi.org/10.1037/emo0000061

Ponari, M., Rodriguez-Cuadrado, S., Vinson, D., Fox, N., Costa, A., & Vigliocco, G. (2014). Processing advantage for emotional words in bilingual speakers. *Emotion*, *15*(January), 644–652. https://doi.org/10.1037/emo0000061

Pratto, F., & John, O. P. (1991). Automatic Vigilance: The Attention-Grabbing Power of Negative Social Information. *Journal of Personality and Social Psychology*, *61*(3), 380–391. https://doi.org/10.1037/0022-3514.61.3.380

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *The Behabioral and Brain Sciences*, *4*, 515–526. https://doi.org/10.1016/j.celrep.2011.1011.1001.7.

Puntoni, S., de Langhe, B., & van Osselaer, S. M. J. (2009). Bilingualism and the Emotional Intensity of Advertising Language. *Journal of Consumer Research*, *35*(6), 1012–1025. https://doi.org/10.1086/595022

Pyszczynski, T., Holt, K., & Greenberg, J. (1987). Depression, self-focused attention, and expectancies for positive and negative future life events for self and others. *Journal of Personality and Social Psychology*, *52*, 994–1001. https://doi.org/10.1037/0022-3514.52.5.994

R Core Team, I. (2017). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna (2014). R Foundation for Statistical Computing Authors R Core Team.

Recio, G., Conrad, M., Hansen, L. B., & Jacobs, A. M. (2014). On pleasure and thrill: The interplay between arousal and valence during visual word recognition. *Brain and Language*, *134*, 34–43. https://doi.org/10.1016/j.bandl.2014.03.009

Redondo, J., Fraga, I., Padrón, I., & Comesaña, M. (2007). The Spanish adaptation of ANEW (affective norms for English words). *Behavior Research Methods*, *39*(3), 600–605. https://doi.org/10.3758/BF03193031

Rickard, N. S. (2004). Intense emotional responses to music: a test of the physiological arousal hypothesis. *Psychology of Music*, *32*(4), 371–388. https://doi.org/10.1177/0305735604046096

Rottenberg, J., & Gotlib, I. (2004). Socioemotional Functioning in Depression. In *Mood Disorders: A Handbook of Science and Practice* (pp. 61–77).

Rottenberg, J., Gross, J. J., & Gotlib, I. H. (2005). Emotion context insensitivity in major depressive disorder. *Journal of Abnormal Psychology*, *114*(4), 627–639. https://doi.org/10.1037/0021-843X.114.4.627

Rottenberg, J., & Hindash, A. C. (2015). Emerging evidence for emotion context insensitivity in depression. *Current Opinion in Psychology*, *4*, 1–5. https://doi.org/10.1016/j.copsyc.2014.12.025

Rush, J., Trivedi, M. H., Ibrahim, H. M., Carmody, T. J., Arnow, B., Klein, D. N., … Keller, M. B. (2003). The 16-Item Quick Inventory of Depressive Symptomatology. *Depression*, *54*(5), 573–583. https://doi.org/10.1016/S0006-3223(03)01866-8

Russell, C. J., & Bobko, P. (1992). Moderated Regression Analysis and Likert Scales Too Coarse for Comfort. *Journal of Applied Psychology*, *77*(3), 336–342.

Russell, J. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*. Retrieved from http://psycnet.apa.org/psycinfo/1981-25062-001

Sandell, K., & Bornäs, H. (2017). Functioning Numbness Instead of Feelings as a Direction: Young Adults' Experiences of Antidepressant Use. *Sociology*, *51*(3), 543–558. https://doi.org/10.1177/0038038515591947

Santiago-Rivera, A. L., & Altarriba, J. (2002). The Role of Language in Therapy With the Spanish-English Bilingual Client. *Professional Psychology: Research and Practice*, *33*(1), 30–38. https://doi.org/10.1037/0735-7028.33.1.30

Schutte, N. S., Malouff, J. M., Bobik, C., Coston, T. D., Greeson, C., Jedlicka, C., … Wendorf, G. (2001). Emotional intelligence and interpersonal relations. *Journal of Social Psychology*, *141*(4), 523–536. https://doi.org/10.1080/00224540109600569

Scott, G., O'Donnell, P., Leuthold, H., & Sereno, S. (2009). Early emotion word processing: evidence from event-related potentials. *Biological Psychology*, *80*(1), 95–104. https://doi.org/S0301-0511(08)00073-2 [pii]\r10.1016/j.biopsycho.2008.03.010

Scott, G., O'Donnell, P., & Sereno, S. (2012). Emotion words affect eye fixations during reading. *Journal of Experimental Psychology: Learning Memory and Cognition*, *38*(3), 783–792. https://doi.org/10.1037/a0027209

Segalowitz, N., Trofimovich, P., Gatbonton, E., & Sokolovskaya, A. (2008). Feeling affect in a second language: The role of word recognition automaticity. *The Mental Lexicon*, *3*(1), 47–71. https://doi.org/10.1075/ml.3.1.05seg

Sharot, T. (2011). The optimism bias. *Current Biology*, *21*(23), R941–R945. https://doi.org/10.1016/j.cub.2011.10.030

Sheikh, N. A., & Titone, D. A. (2013). Sensorimotor and linguistic information attenuate emotional word processing benefits: An eye-movement study. *Emotion, 13*(6), 1107–1121. https://doi.org/10.1037/a0032417

Sheikh, N., & Titone, D. (2015). The embodiment of emotional words in a second language: An eye-movement study. *Cognition & Emotion,* (April), 1–13. https://doi.org/10.1080/02699931.2015.1018144

Shields, M., Gorber, S. C., Janssen, I., & Tremblay, M. S. (2011). Bias in self-reported estimates of obesity in Canadian health surveys: An update on correction equations for adults. *Health Reports*, *22*(3).

Sigmon, S. T., & Nelson-Gray, R. O. (1992). Sensitivity to Aversive Events in Depression: Antecedent, Concomitant, or Consequent? *Journal of Psychopathology and Behavioral Assessment*, *14*(3), 225–246.

Skehan, P. (1991). Individual Differences in Second Language Learning. *Studies in Second Language Acquisition*, *13*(02), 275. https://doi.org/10.1017/S0272263100009979

Sloan, D. M., Strauss, M. E., Quirk, S. W., & Sajatovic, M. (1997). Subjective and expressive emotional responses in depression. *Journal of Affective Disorders*, *46*, 135–141. https://doi.org/https://doi.org/10.1016/S0165-0327(97)00097-9

Sloan, D. M., Strauss, M. E., & Wisner, K. L. (2001). Diminished Response to Pleasant Stimuli by Depressed Women. *Journal of Abnormal Psychology*, *110*(3), 488–493. https://doi.org/I0.1037//0021-843X.110.3.488

Smith, D. P., Hillman, C. H., & Duley, A. R. (2005). Influences of Age on Emotional Reactivity During Picture Processing. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *60*(1), P49–P56. https://doi.org/10.1093/geronb/60.1.P49

Soares, A. P., Comesaña, M., Pinheiro, A. P., Simões, A., & Frade, S. (2012). Behavior and Research Methods Corresponding author : *Behavior Research Methods*, *44*(1), 256–269. https://doi.org/10.3758/s13428-011-0131-7

Stadthagen-González, H., Ferré, P., Pérez-Sánchez, M. A., Imbault, C., & Hinojosa, J. A. (2017). Norms for 10,491 Spanish words for five discrete emotions: Happiness, disgust, anger, fear, and sadness. *Behavior Research Methods*. https://doi.org/10.3758/s13428-017-0962-y

Stadthagen-Gonzalez, H., Imbault, C., Pérez Sánchez, M. A., & Brysbaert, M. (2017). Norms of valence and arousal for 14,031 Spanish words. *Behavior Research Methods*, 1–45. https://doi.org/10.3758/s13428-015-0700-2

Statistics Canada. (2013). *National Household Survey*. *Immigration and Ethnocultural Diversity in Canada* (Vol. 99-010-X20). https://doi.org/99-010-X2011001

Sutton, T. M., Altarriba, J., Gianico, J. L., & Basnight-Brown, D. M. (2007). The automatic access of emotion: Emotional Stroop effects in Spanish-English bilingual speakers. *Cognition and Emotion*, *21*(5), 1077–1090. https://doi.org/10.1080/02699930601054133

Thoma, P., Norra, C., Juckel, G., Suchan, B., & Bellebaum, C. (2015). Performance monitoring and empathy during active and observational learning in patients with major depression. *Biological Psychology*, *109*, 222–231. https://doi.org/10.1016/j.biopsycho.2015.06.002

Tsai, J. L., Knutson, B., & Fung, H. H. (2006). Cultural variation in affect valuation. *Journal of Personality and Social Psychology*, *90*(2), 288–307. https://doi.org/10.1037/0022-3514.90.2.288

Velez-Uribe, I., & Rosselli, M. (2019). The auditory and visual appraisal of emotion-related words in Spanish–English bilinguals. *Bilingualism: Language and Cognition*, *22*(1), 30–46. https://doi.org/10.1017/S1366728917000517

Vinson, D., Ponari, M., & Vigliocco, G. (2014). How does emotional content affect lexical processing? *Cognition & Emotion*, *28*(4), 737–746. https://doi.org/10.1080/02699931.2013.851068

Võ, M. L. H., Jacobs, A. M., Kuchinke, L., Hofmann, M., Conrad, M., Schacht, A., & Hutzler, F. (2008). The coupling of emotion and cognition in the eye: Introducing the pupil old/new

effect. *Psychophysiology*, *45*(1), 130–140. https://doi.org/10.1111/j.1469-8986.2007.00606.x

Wang, Y. guang, Wang, Y. qiang, Chen, S. lin, Zhu, C. yan, & Wang, K. (2008). Theory of mind disability in major depression with or without psychotic symptoms: A componential view. *Psychiatry Research*, *161*(2), 153–161. https://doi.org/10.1016/j.psychres.2007.07.018

Waris, O., Soveri, A., Ahti, M., Hoffing, R. C., Ventus, D., Jaeggi, S. M., … Laine, M. (2017). A latent factor analysis of working memory measures using large-scale data. *Frontiers in Psychology*, *8*(JUN), 1–14. https://doi.org/10.3389/fpsyg.2017.01062

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191–1207. https://doi.org/10.3758/s13428-012-0314-x

Warriner, A. B., Shore, D. I., Schmidt, L. A., Imbault, C. L., & Kuperman, V. (2017). Sliding into happiness: A new tool for measuring affective responses to words. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *71*(1), 71–88. https://doi.org/10.1037/cep0000112

Watson, D., Clark, L. A., & Carey, G. (1988). Positive and negative affectivity and their relation to anxiety and depressive disorders. *Journal of Abnormal Psychology*, *97*(3), 346–353. https://doi.org/10.1037/0021-843X.97.3.346

Weisleder, A., & Fernald, A. (2013). Talking to Children Matters: Early Language Experience Strengthens Processing and Builds Vocabulary. *Psychological Science*, *24*(11), 2143–2152. https://doi.org/10.1177/0956797613488145

Westbury, C., Keith, J., Briesemeister, B. B., Hofmann, M. J., & Jacobs, A. M. (2015). Avoid violence, rioting, and outrage; approach celebration, delight, and strength: Using large text corpora to compute valence, arousal, and the basic emotions. *The Quarterly Journal of Experimental Psychology*, *0*(0), 1–24. https://doi.org/10.1080/17470218.2014.970204

Wickham, H., & Chang, W. (2016). Create elegant data visualisations using the grammar of graphics. https://doi.org/10.1093/bioinformatics/btr406

Williams, J. M., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and

psychopathology. *Psychological Bulletin*, *120*(1), 3–24. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8711015

Winskel, H. (2013). The emotional Stroop task and emotionality rating of negative and neutral words in late Thai-English bilinguals. *International Journal of Psychology*. Taylor & Francis. https://doi.org/10.1080/00207594.2013.793800

Wolkenstein, L., Schönenberg, M., Schirm, E., & Hautzinger, M. (2011). I can see what you feel, but i can’t deal with it: Impaired theory of mind in depression. *Journal of Affective Disorders*, *132*(1–2), 104–111. https://doi.org/10.1016/j.jad.2011.02.010

Wolter, B. (2001). COMPARING THE L1 AND L2 MENTAL LEXICON A Depth of Individual Word Knowledge Model. *Studies in Second Language Acquisition*, *23*(1), S0272263101001024. https://doi.org/10.1017/S0272263101001024

Zuroff, D., Koestner, R., Moskowitz, D. S., McBride, C., Marshall, M., & Bagby, M. (2007). Autonomous motivation for therapy: A new common factor in brief treatments for depression. *Psychotherapy Research*, *17*(2), 137–148. https://doi.org/10.1080/10503300600919380

Author Contributions

CI and VK designed the study and carried out the initial acquisition, analysis, interpretation of data drafting and revising the manuscript and have given final approval of the version to be published.

Competing Financial Interests

The authors declare no competing interests.

Chapter 5

5.0 General

The goal of this thesis was to investigate group differences in emotional processing, specifically with linguistic stimuli. In this thesis, I used valence as our main dimension of emotionality. I used two methods to measure valence, a typical ratings scale (from 1 (negative) to 9 (positive)) and the sliding scale, which was recently proposed by Warriner et al. (2018). The measurement of emotional processing is at the cognitive level (summarized in Kleinginna & Kleinginna (1981)) – I specifically asked my participants to appraise how the stimulus makes them feel. I measured people's overt feelings (and thoughts) towards my stimuli, not their visceral, physiological reactions.

So far, the field of individual and group differences on emotional processing has mainly focused on the neural activation of emotional stimuli, and has looked at individual differences at the personality level. In this thesis, I used two different methods to measure emotional processing, and I investigated the differences in emotional reactivity of two groups that were not defined at the personality level; those with and without depressive symptoms, and native and non-native speakers of English. These two groups both have very contested findings as to their emotional reactivity, and my goal was to further investigate how these two groups process the emotion of linguistic stimuli.

It is incredibly interesting that we are able to see such large effects in our measurement of emotional perception. As discussed in the Introduction, emotional processing and perception is an automatic, evolutionarily guided process. We all have similar physiological circuitry that responds to affective stimuli, and yet large differences can emerge from different groups. The differences between those with and without depressive symptoms may be explained at the level of physiology and hormones, but the differences we observed between native and non-native speakers has to do with how these groups interact with their language. The differences we observed originate purely from how these groups learned words and how they use and interact with their language and the world around them.

5.1 Summary of findings

Chapter 2 investigated how and why non-native speakers process the emotion of words differently than native speakers. I found that overall, non-native speakers had attenuated responses to emotional words at the extreme ends of the scale – words rated as either extremely positive or negative by native speakers were rated more neutrally by non-native speakers of English. This supports many theories that suggest that non-native speakers have reduced emotional processing in their non-native language (Anooshian & Hertel, 1994; Caldwell-Harris et al., 2011; J. M. Dewaele & Nakano, 2013; Harris et al., 2003; Santiago-Rivera & Altarriba, 2002; Sutton et al., 2007). Furthermore, I found that this attenuation is exacerbated in contexts where non-native speakers tend to not use English. Contexts where they would use English in Canada, such as in the workplace or at school, shows similar levels of emotional reactivity as native speakers of English. On the other hand, contexts where non-native speakers are more likely to use their native language, such as when using swear or taboo words, shows greater levels of attenuation. This supports theories that the emotion of a word is learned through and together with sensorimotor interactions using that word (Barsalou, 1999; Kousta et al., 2011). Furthermore, both length of time spent living in Canada as well as English reading proficiency were predictors of how native-like emotion responses to words were. Again, this supports the theory that the emotional connotation is learned through everyday interactions with a word, not simply by learning the denotation of the word (perhaps through a book or in a classroom).

Chapter 3 fills the void for a useful method that allows for the fine-grained measure of explicit behavioural affective responses to words. The method used in Chapter 2 requires participants to indicate their emotional response along a 1-9 point rating scale. Although this method has been used in many studies in the past, it has its drawbacks. The first, and most critical point when it comes to studying subtle differences, is how coarse-grained the method is. Nine separate points is not enough to be able to measure subtle individual differences. Since emotion is an evolved response, one that we all have similar physiology for, the differences in how we perceive emotion will be small. But even those these differences are small, they can still have interesting and important effects. If we think about another group difference, such as height, we see that small differences are actually noticeable and impact our lives. For example, the average height of a male in Canada is 175.8cm, and the average for a female is 162.1cm (Shields, Gorber, Janssen, & Tremblay, 2011). This is only a 7.8% difference. This is a small difference, yet we all accept that males are, on average, taller than females. If we expand this difference to traditional valence ratings (out of 9), the difference between men and women might average out to be 0.7 points on a scale out of 9. This is not a large enough difference to be able to consistently measure on such a coarse-grained scale. On the other hand, in this chapter, we tested a sliding scale method that allows participants to give an emotional response to a word along a 648-point continuum. A difference of 7.8% results in a 50.5-pixel difference, which is easily measured with this finer-grained scale. This demonstrates how the sliding scale allows for subtle differences between individuals to be measured.

The second issue with a rating scale is that it collects ordinal data, but is interpreted as interval data. The sliding scale collects interval data that is more intuitively interpretable. I used 3 different methods to evaluate the reliability of the sliding scale. In all three methods, I found mid to high levels of reliability. Reliability is crucial for the study of individual differences. If a method produces different results within a single person, any differences across people cannot be attributed to individual differences, but rather problems in the method of measurement.

Chapter 4 uses the sliding scale method to measure the affective differences between those with and without depressive symptoms. There are several conflicting theories about the emotional reactivity of those with depression. Some propose that depression is the absence of positivity, others propose that it is the potentiation of negativity, and other theories suggest that depression is the absence of both positivity and negativity. In Chapter 4, I explicitly asked those with depressive symptoms to give their emotional responses to words. I compared these responses with people who do not have depressive symptoms. I found that those with depressive symptoms experience emotional context insensitivity; their emotional responses were dampened in both the positive and negative domain.

I also investigated whether these two groups were aware of the emotional differences between them. Given that researchers of depression argue about the emotional reactivity of people with depression, I suspected that the general population would be unable to accurately predict the emotional reactivity of depression. Surprisingly, I found that those without depression accurately predicted the attenuated emotional responses of those with depressive symptoms. This understanding is beneficial to the prognosis of depression. Many studies suggest that the social network surrounding someone with depression has a large impact on the remission rates and severity of depression (George et al., 1989; Peirce et al., 2000). Another interesting finding was how those with depressive symptoms predicted how individuals without depression felt. I observed that their emotional responses did not change with the perspective taking instruction – they responded with the same (attenuated) emotional reactivity when responding for someone else as they did for themselves. This leads us to believe that those with depression are unable to understand what is it like to not be depressed. This finding is supported by the hopelessness theory of depression. This theory posits that one of the causes of depression is the belief that desired outcomes will not occur, and aversive outcomes are most likely (Abramson et al., 1989). Our participants with depressive symptoms could not predict desired, non-depressed outcomes. These beliefs are also detrimental to the prognosis of depression; autonomous motivation, or the motivation "to get better" is associated with lessened depressive symptomatology (Zuroff et al., 2007). The missing motivation is the ability to understand and feel what it is like to not feel depressed.

5.2 Limitations of findings

One of the greatest limitations for the entirety of this thesis is that I asked my participants to give their cognitive emotional appraisal of my stimuli. I did not measure their physiological responses, so I do not know whether the participant actually feels their reported emotion, or if they semantically understand that the word is positive or negative and give the resulting rating. My goal is to measure the former – actual emotional responses to words – but it has been debated that ratings studies actually measure the latter. This question was debated in a recent paper by Itkes, Eviatar, & Kron (2019). They argued that there are two different ways to measure valence; activating the affective response profile (and actually feeling the emotional response) and activating the semantic representation of the word (knowing that the particular word is positive or negative). These two dimensions are similar to the affective and cognitive definitions given by Kleinginna & Kleinginna (1981). Itkes, Eviatar & Kron found that these two measures are distinct, and that they can be captured separately through different instructions ("rate the valence of the stimulus" vs "rate the valence of your feeling"; Itkes, Kimchi, Kron, & Carmel, 2017). In my studies, I may be only activating the semantic representation of the word, rather than the affective response profile. We are unable to determine whether the words presented to the participants actually influenced their affect, or if the participants were solely responding to the meaning of the word.

Another limitation is the type of stimuli I used in these studies. Although we discussed that linguistic stimuli produce similar emotional responses to other types of stimuli, it is still a less ecologically valid medium than pictures or videos. This may also lead back to the first limitation, more ecologically valid stimuli may induce more visceral, affective responses, whereas less ecologically valid stimuli may produce more semantic response. Future studies can investigate whether we find similar findings with pictorial stimuli. It would be interesting to see whether our findings are intensified by more valid stimuli, and whether the differences between populations become even larger.

Another limitation in my experiments is shared by many psychological studies. Our participants are very homogenous. Chapters 2 and 3 used an undergraduate population sample. In doing so, our participants are of similar ages (18-25), similar level of education, and relatively high socioeconomic status. These invariable demographics are groups that we should study to investigate if they differ in their emotional processing (e.g., young compared to old). Furthermore, groups with larger variance are the perfect example of a study where an individual differences approach would be useful. This idea is stated perfectly by Jonathan Haidt (who was paraphrasing Robert Zajonc) that "cognitive psychology is social psychology with all the interesting variables set to zero". The variability between participants, which is so often ignored or intentionally minimized, is exactly what this thesis studies and what I find interesting.

Working with a population with depressive symptoms comes with its own limitations. Chapter 4 has some limitations that are specific to working with a special population. The first is the procedure and instructions that we used. I did not instruct my participants to first as themselves, and then respond as someone with/without depression. My point in doing so was deliberate; I did not want to explicitly identify my participants as depressed or not depressed. Identifying participants within their groups can cause them to respond differently due to cognitive and social response bias (Orne, 1962). This could have caused my participants with depressive symptoms to respond as a stereotyped "depressed" individual, rather than themselves. To remedy this situation, in future studies I would start each experiment with a block where participants respond as themselves, and a second block with the perspective taking instruction is introduced. In doing so, I would capture the actual emotional reactivity of each participant I am planning to study without explicitly identifying them within their group.

Another possible limitation of Chapter 4 is my lack of knowledge about the participants' medical history. I do not have any knowledge about the possible medications the participants are taking, which could lead to differences in their emotional reactivity. Some antidepressants can lead "emotional numbness", the lack of emotional feeling (Cartwright, Gibson, Read, Cowan, & Dehar, 2016; Sandell & Bornäs, 2017). This symptom is exactly what I measured in our study (overall numbness across the entire emotion spectrum), and because I am unaware of my participants' medical history, I cannot say whether my observations are due to their depressive symptoms, or their potential antidepressant use. In future studies on depression, it is pertinent to ask participants about their medical background and medications.

Furthermore, I tested depressive symptoms in a screening questionnaire a month before releasing this experiment. Since I do not know the clinical status of our participants, I do not know if their depressive symptoms were a short-lived emotion (ie. my participants felt depressed in the moment, but do not have long lasting depression) or if their symptoms were due to long-lasting depressed mood (which would be a sign of clinical depression). This means that any conclusions I make about these findings can only be made for people with *depressive symptoms* and not people with depression.

5.3 Implications and Significance

This thesis provides motivation to other emotion researchers to include individual differences in all levels of analysis. This thesis has shown that there are vast differences between groups that may have once been clustered together. Such differences may have resulted in null results in the past (Canli, Sivers, Whitfield, Gotlib, & Gabrieli, 2002), but, as I have shown, can uncover interesting and novel findings.

As I have discussed, emotion is an evolutionary response to guide us from potentially dangerous situations. We are all hard-wired in similar ways, and yet, as I have shown, we all responded slightly differently to stimuli. Some of these differences can be due to different levels of hormones (ex. those with and without depressive symptoms), but others are due to how we have interacted with our environment (ex. native and non-native speakers of English). There are many more groups that deserve to be studied to uncover any hidden emotional processing differences between them.

Firstly, this thesis presented the measure of emotional reactivity for two groups. The emotional reactivity of both groups has been hotly debated in the literature, with different groups of researchers arguing for opposite types of emotional reactivity. In the second language literature, some studies find that non-native speakers of a language process the emotion of their language in a similar fashion to native speakers (Ayçiçeği-Dinn & Caldwell-Harris, 2009; Eilola et al., 2007; Ferré et al., 2010; Ponari, Rodriguez-Cuadrado, Vinson, et al., 2014). On the other hand, some studies find completely conflicting results, saying that non-native speakers have dampened emotional responses at the extreme ends of the valence scale (Anooshian & Hertel, 1994; Caldwell-Harris et al., 2011; J. M. Dewaele & Nakano, 2013; Harris et al., 2003; Santiago-Rivera & Altarriba, 2002; Sutton et al., 2007). Chapter 2 provided a large dataset worth of valence and arousal ratings to help disentangle these questions. I found that non-native speakers do indeed experience dampening at extreme ends of the scale. Furthermore, I used a novel approach in this literature by separating our stimulus list into several different lists representing different semantic categories (Pennebaker et al., 2007). In doing so, I was able to pinpoint where and why non-native speakers differ from native speakers of a language. I hope that other researchers follow this approach and potentially take a closer look at their stimulus lists, whether it be in bilingualism research, or psychological research in general. Additionally, Chapter 2 provides the largest dataset of affective norms collected from non-native speakers. Researchers should use these norms, along with traditional native-speaker norms, in their research about second language learners.

A more practical implication of Chapter 2 is its ability to inform teachers of English as a second language. I observed that the way a non-native speaker uses language affects how native-like their emotionality is towards that language. I found that using a non-native language and encountering sensorimotor experiences with the language in their everyday life causes them to become more emotional about that language. Teachers of English as a second language can teach English in a more representative way – have students use different semantic domains of English in their conversations to allow for more emotional grounding.

The second group that I chose to study was those with depressive symptoms. Despite being a hot topic to research, there has been a lot of debate about the emotional reactivity of those with depression. Similar to non-native speakers of English, researchers propose opposite and conflicting findings about the emotional reactivity of those with depression. Some researchers propose that depression is a fundamental dampening of positive emotion (Lewinsohn et al., 1973; Sloan et al., 1997, 2001), and a strong intensification of negative emotions (Allen et al., 1999; Golin et al., 1977; Sigmon & Nelson-Gray, 1992). Other researchers find that those with depression have emotion context insensitivity – dampening on both ends of the scale (Rottenberg et al., 2005; Rottenberg & Hindash, 2015). In my study, I explicitly tested the emotional responses of those with depressive symptoms to uncover the true emotional reactivity of this population. I found overall dampening of emotion, supporting theories of emotion context insensitivity. Furthermore, this study went above testing emotional reactivity of depression and instead delved into the thoughts and beliefs those with and without depressive symptoms have of each other. I found that those without depression can surprisingly predict the emotional reactivity of those with depressive symptoms. This provides uplifting evidence for those with depression, hopefully supporting their recovery and providing a feeling of understanding. On the other hand, I also tested if those with depressive symptoms can imagine what it feels like to not have depression. What I found was surprising – that they couldn't take on the perspective of someone without depression. These findings can help inform why depression ends up being a spiral for some, and can inform clinicians about possible ways of promoting remission.

This thesis also provides evidence for why we should be using a finer-grained approach to measuring valence. I strongly believe that countless subtle individual differences have been lost throughout the emotion literature due to too coarse of measurement. Even if researchers investigate emotion with an individual difference viewpoint, they may be unable to find these differences using a rating-style scale that is typically used in the literature.

Overall, the goal of this thesis was to provide motivation for future researchers to use an individual and group differences approach in their research, and to see individual variation as an interesting result and not nuisance noise.

5.4 Future Directions

There is a wealth of other group and individual differences that I believe can be investigated and that would align with this thesis. They all start with breaking down our sample population into smaller groups to investigate the participant's own variables, and their effect on emotional reactivity.

Larsen & Ketelaar (1991) found that an individual's Big Five traits affected how emotionally susceptible they are to mood inductions; those higher in neuroticism showed increased emotional reactivity to negative-mood induction, and those high in extraversion greater emotional reactivity in response to positive-mood induction. Using the slider, we can have a fine-grained observation of any differences between those high and low in several personality characteristics. Warriner et al. (2018) have already given evidence for the effect of shyness on one's emotion perception. We can extend these findings by investigating the effect of the Big Five personality traits on an individual's own emotion perception and emotional reactivity.

Another possible direction to study is how one's environment affects the way they process emotion. Kuppens et al. (2017) showed that cultural background can affect the way individuals describe their own experience. They stressed the importance of studying cultures other than typical Western cultures that are used in the bulk of Psychology. They proposed that Western and Eastern cultures are very different, and that many conclusions from Western studies cannot be generalized to non-Western populations. In Western societies, Tsai, Knutson, & Fung (2006) showed that people preferred high arousing, positive states, but Eastern societies prefer low arousing, positive states. They suggested that this is due to the interpersonal goals of each culture. Using the slider paradigm to further explore these differences, we may be able to uncover even more differences between the two cultures. As well, we can study groups from this thesis from an Eastern society to see if my findings hold.

Another interesting group to investigate is how young and old adults differ in how they process emotion. If there are any differences between these two groups, we can use the perspective taking abilities of the slider task to see whether these two groups have the ability to understand how the other group responds. There is some evidence to believe that young and older adults process emotion differently; older adults have a greater startle blink response to negative stimuli than young adults do, suggesting that they may be more affected by negative valence (Smith, Hillman, & Duley, 2005). Going along with the theme of this thesis, this topic is contested. Other researchers find conflicting results, stating that there is no difference in the way the young and old adults process emotion (Levenson, Carstensen, Friesen, & Ekman, 1991). It would be interesting to see how these differences would play out in the slider paradigm, and if there would be an overall shift in emotional reactivity with age. Furthermore, it would be interesting to see if there are differences in how young and older adults measure the emotion of specific semantic categories. The majority of studies in psychology focus on younger adults, and have stimulus lists catered to younger adults. Indeed, Kunzmann & Grühn (2005) found greater emotional reactions to stimuli that are thematically relevant to older adults; above what is to be expected based on previous literature.

Going along with this, separating our stimulus lists into specific semantic categories, similarly to how we did in Chapter 2, would be an interesting way to investigate individual differences. Warriner et al. (2018) found inconsistent results in how those higher in shyness tended to be further away from all stimuli. If a more thematically relevant stimulus list was used (perhaps a list with social connotations), we could find more consistent results.

5.5 Conclusions

Overall, this thesis has stressed the importance of using an individual and groups differences approach when designing and analyzing experiments. Future researchers should be motivated to split their participants based on their individual characteristics to uncover new and interesting findings that are often lost at the aggregate level of analysis.

References

Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review*, *96*(2), 358–372. https://doi.org/10.1037/0033-295X.96.2.358

Albaum, G., Best, R., & Hawkins, D. E. L. (1981). Continuous vs Discrete Semantic Differential Rating Scales. *Psychological Reports*, *49*, 83–86.

Allen, N. B., Trinder, J., & Brennan, C. (1999). Affective Startle Modulation in Clinical Depression: Preliminary Findings. *Biological Psychiatry*, *46*(4), 542–550. https://doi.org/https://doi.org/10.1016/S0006-3223(99)00025-6

Altarriba, J., & Bauer, L. M. (2004). The Distinctiveness of Emotion Concepts: A Comparison between Emotion, Abstract, and Concrete Words. *The American Journal of Psychology*, *117*(3), 389–410.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (DSM-5®). *American Psychiatric Pub*.

Anooshian, L. J., & Hertel, P. T. (1994). Emotionality in free recall: Language specificity in bilingual memory. *Cognition and Emotion*, *8*(6), 503–514. https://doi.org/10.1080/02699939408408956

Ayçiçeği-Dinn, A., & Caldwell-Harris, C. L. (2009). Emotion-memory effects in bilingual speakers: A levels-of-processing approach*. *Bilingualism: Language and Cognition*, *12*(03), 291. https://doi.org/10.1017/S1366728909990125

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(04), 577–660. https://doi.org/10.1017/S0140525X99002149

Batson, C. D., Early, S., & Salvarani, G. (1997). Perspective Taking: Imagining How Another Feels Versus Imaging How You Would Feel. *Personality and Social Psychology Bulletin*. https://doi.org/10.1177/0146167297237008

Bayer, M., Sommer, W., & Schacht, A. (2011). Emotional words impact the mind but not the body: Evidence from pupillary responses. *Psychophysiology*, *48*(11), 1554–1562. https://doi.org/10.1111/j.1469-8986.2011.01219.x

Betella, A., & Verschure, P. F. M. J. (2016). The affective slider: A digital self-assessment scale for the measurement of human emotions. *PLoS ONE*, *11*(2), 1–11. https://doi.org/10.1371/journal.pone.0148037

Bonin, P., Aubert, L., Malardier, N., & Niedenthal, P. M. (2016). Subjective Et De Valence Émotionnelle Pour 866 Mots, *103*, 655–694.

Bradley, M. M., & Lang, P. P. J. (1999). Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings. *Psychology*, *Technical*(C–1), 0. https://doi.org/10.1109/MIC.2008.114

Brase, J., & Mani, N. (2017). Effects of learning context on the acquisition and processing of emotional words in bilinguals. *Emotion*, *17*(4), 628–639. https://doi.org/10.1037/emo0000263

Briesemeister, B. B., Kuchinke, L., & Jacobs, A. M. (2011). Discrete emotion effects on lexical decision response times. *PLoS ONE*, *6*(8). https://doi.org/10.1371/journal.pone.0023743

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990. https://doi.org/10.3758/BRM.41.4.977

Bylsma, L. M., Morris, B. H., & Rottenberg, J. (2008). A meta-analysis of emotional reactivity in major depressive disorder. *Clinical Psychology Review*, *28*(4), 676–691. https://doi.org/10.1016/j.cpr.2007.10.001

Caldwell-Harris, C. L. (2015). Emotionality Differences Between a Native and Foreign Language. *Current Directions in Psychological Science*, *24*(3), 214–219. https://doi.org/10.1177/0963721414566268

Caldwell-Harris, C. L., Tong, J., Lung, W., & Poo, S. (2011). Physiological reactivity to emotional phrases in Mandarin--English bilinguals. *International Journal of Bilingualism*, *15*(3), 329–352. https://doi.org/10.1177/1367006910379262

Canli, T., Sivers, H., Whitfield, S. L., Gotlib, I. H., & Gabrieli, J. D. E. (2002). Amygdala response to happy faces as a function of extraversion. *Science*, *296*(5576), 2191. https://doi.org/10.1126/science.1068749

Cartwright, C., Gibson, K., Read, J., Cowan, O., & Dehar, T. (2016). Long-term antidepressant use: patient perspectives of benefits and adverse effects. *Patient Preference and Adherence*, *Volume 10*, 1401–1407. https://doi.org/10.2147/PPA.S110632

Carver, C. S., & White, T. L. (1994). Behavioral Inhibition, Behavioral Activation and affective responses to impending reward and punishment. *Journal of Personality and Social Psychology*, *67*, 319–333.

Chen, M., & Bargh, J. A. (1997). Nonconscious behavioral confirmation processes: The self-fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology*, *33*(5), 541–560. https://doi.org/10.1006/jesp.1997.1329

Citron, F. M. M., Weekes, B. S., & Ferstl, E. C. (2013). Effects of valence and arousal on written word recognition: Time course and ERP correlates. *Neuroscience Letters*, *533*, 90–95.

Clark, J., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*(3), 149–210. https://doi.org/10.1007/BF01320076

Cutrona, C. E., & Troutman, B. R. (1986). Social Support , Infant Temperament , and Parenting Self-Efficacy : A Mediational Model of Postpartum Depression Author (s): Carolyn E . Cutrona and Beth R . Troutman Published by : Wiley on behalf of the Society for Research in Child Development Stable, *57*(6), 1507–1518.

Daneman, M., & Carpenter, P. A. (1980). Individual Differences in Working Memory and Reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466.

De Grauwe, S., Willems, R. M., Rueschemeyer, S. A., Lemhöfer, K., & Schriefers, H. (2014). Embodied language in first- and second-language speakers: Neural correlates of processing motor verbs. *Neuropsychologia*, *56*(1), 334–349. https://doi.org/10.1016/j.neuropsychologia.2014.02.003

De Pascalis, V., Fiore, A. D., & Sparita, A. (1996). Personality, event-related potential (ERP) and heart rate (HR): An investigation of Gray's theory. *Personality and Individual Differences*, *20*(6), 733–746. https://doi.org/10.1016/0191-8869(96)00016-5

De Pascalis, V., Strippoli, E., Riccardi, P., & Vergari, F. (2004). Personality, event-related potential (ERP) and heart rate (HR) in emotional word processing. *Personality and Individual Differences*, *36*(4), 873–891. https://doi.org/10.1016/S0191-8869(03)00159-4

Decety, J., & Batson, C. D. (2007). Social neuroscience approaches to interpersonal sensitivity. *Social Neuroscience*, *2*(3–4), 151–157. https://doi.org/10.1080/17470910701506060

Degner, J., Doycheva, C., & Wentura, D. (2012). It matters how much you talk: On the automaticity of affective connotations of first and second language words. *Bilingualism*, *15*(1), 181–189. https://doi.org/10.1017/S1366728911000095

Dewaele, J.-M. (2004). The Emotional Force of Swearwords and Taboo Words in the Speech of Multilinguals. *Journal of Multilingual and Multicultural Development*, *25*(23), 204–222. https://doi.org/10.1080/01434630408666529

Dewaele, J. M., & Nakano, S. (2013). Multilinguals' perceptions of feeling different when switching languages. *Journal of Multilingual and Multicultural Development*, *34*(2), 107–120. https://doi.org/10.1080/01434632.2012.712133

Dresler, T., Mériau, K., Heekeren, H. R., & Van Der Meer, E. (2009). Emotional Stroop task: Effect of word arousal and subject anxiety on emotional interference. *Psychological Research*, *73*(3), 364–371. https://doi.org/10.1007/s00426-008-0154-6

Ehrman, M. E., Leaver, B. Lou, & Oxford, R. L. (2003). A brief overview of individual differences in second language learning. *System*, *31*(3), 313–330. https://doi.org/10.1016/S0346-251X(03)00045-9

Eilola, T. M., Havelka, J., & Sharma, D. (2007). Emotional activation in the first and second language. *Cognition and Emotion*, *21*(5), 1064–1076. https://doi.org/10.1080/02699930601054109

Eisman, A. B., Stoddard, S. a, Heinze, J., Caldwell, C. H., & Zimmerman, M. a. (2015). Depressive symptoms, social support, and violence exposure among urban youth: A longitudinal study of resilience. *Developmental Psychology*, *51*(9), 1307–1316. https://doi.org/10.1037/a0039501

Ekman, P. (1992). An Argument for Basic Emotions. *Cognition and Emotion*, *6*(3–4), 169–200. https://doi.org/10.1080/02699939208411068

Elliot, A. J. (2006). The Hierarchical Model of Approach-Avoidance Motivation. *Motivation and Emotion*, *30*(2), 111–116. https://doi.org/10.1007/s11031-006-9028-7

Epstein, J., Pan, H., Kocsis, J., Yang, Y., Butler, T., Chusid, J., … Silbersweig, D. (2006). Lack of Ventral Striatal Response to Positive Stimuli inDepressed Versus Normal Subjects. *The American Journal of Psychiatry*, 1784. https://doi.org/10.1176/ajp.2006.163.10.1784

Farrow, T. F. D., & Woodruff, P. W. R. (2008). Empathy in Mental Illness. *Psychiatric Services*, *59*, 581–581. https://doi.org/10.1176/appi.ps.59.5.581

Feldman Barrett, L., & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology*, *74*(4), 967–984. https://doi.org/10.1037/0022-3514.74.4.967

Ferré, P., García, T., Fraga, I., Sánchez-Casas, R., & Molero, M. (2010). Memory for emotional words in bilinguals: Do words have the same emotional intensity in the first and in the second language? *Cognition and Emotion*, *24*(5), 760–785. https://doi.org/10.1080/02699930902985779

Foroni, F. (2015). Do we embody second language? Evidence for ‘partial’ simulation during processing of a second language. *Brain and Cognition*, *99*, 8–16. https://doi.org/10.1016/j.bandc.2015.06.006

George, L., Blazer, D., Hughes, D., & Fowler, N. (1989). Social support and the outcome of major depression. *British Journal of Psychiatry*, *154*, 478–485.

Golin, S., Hartman, S. A., Klatt, E. N., Munz, K., & Wolfgang, G. L. (1977). Effects of Self-esteem Manipulation on Arousal and Reactions to Sad Models in Depressed and Nondepressed College Students. *Journal of Abnormal Psychology*, *86*(4), 435–439.

Haaga, D. A. F., & Beck, A. T. (1995). Perspectives on depressive realism: Implications for cognitive theory of depression. *Behaviour Research and Therapy*, *33*(1), 41–48. https://doi.org/10.1016/0005-7967(94)E0016-C

Harris, C. L., Ayçiçeği, A., & Gleason, J. B. (2003). Taboo words and reprimands elicit greater autonomic reactivity in a first language than in a second language. *Applied Psycholinguistics*, *24*(4), 561–579. https://doi.org/10.1017/S0142716403000286

Harris, C. L., Gleason, J. B., & Ayçiçeği, A. (2006). When is a First Language More Emotional? Psychophysiological Evidence from Bilingual Speakers. In *Bilingual minds: Emotional experience, expression and representation* (pp. 257–283). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.69.5830%0Ahttp://www.zeit.de/wirtschaft/2014-08/russland-import-sanktionen%0Ahttp://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.69.5830%0Ahttp://www.zeit.de/wirtschaft/2014-08/russland-import-sa

Hellman, A. B. (2011). Vocabulary size and depth of word knowledge in adult-onset second language acquisition. *International Journal of Applied Linguistics*, *21*(2), 162–182. https://doi.org/10.1111/j.1473-4192.2010.00265.x

Hinojosa, J. a., Martínez-García, N., Villalba-García, C., Fernández-Folgueiras, U., Sánchez-Carmona, a., Pozo, M. a., & Montoro, P. R. (2015). Affective norms of 875 Spanish words for five discrete emotional categories and two emotional dimensions. *Behavior Research Methods*, (August). https://doi.org/10.3758/s13428-015-0572-5

Imbault, C., Shore, D., & Kuperman, V. (2017). Reliability of the sliding scale for collecting affective responses to words. *Behavior Research Methods*.

Inoue, Y., Yamada, K., & Kanba, S. (2006). Deficit in theory of mind is a risk for relapse of major depression. *Journal of Affective Disorders*, *95*(1–3), 125–127. https://doi.org/10.1016/j.jad.2006.04.018

Inoue, Yumiko, Tonooka, Y., Yamada, K., & Kanba, S. (2004). Deficiency of theory of mind in patients with remitted mood disorder. *Journal of Affective Disorders*, *82*(3), 403–409. https://doi.org/10.1016/j.jad.2004.04.004

Itkes, O., Eviatar, Z., & Kron, A. (2019). Semantic and affective manifestations of ambi (valence). *Cognition and Emotion*, *9931*. https://doi.org/10.1080/02699931.2018.1564249

Itkes, O., Kimchi, R., Kron, A., & Carmel, M. (2017). Dissociating Affective and Semantic Valence, *146*(7), 1–36. https://doi.org/10.1037/xge0000291

Itzhak, I., Vingron, N., Baum, S. R., & Titone, D. (2017). Bilingualism in the real world: How proficiency, emotion, and personality in a second language impact communication in clinical and legal settings. *Translational Issues in Psychological Science*, *3*(1), 48–65. https://doi.org/10.1037/tps0000103

Johnson-Laird, P. N., & Oatley, K. (1989). The Language of Emotions: An Analysis of a Semantic Field. *Cognition and Emotion*, *3*(2), 81–123. https://doi.org/10.1080/02699938908408075

Jolliffe, D., & Farrington, D. P. (2006). Development and validation of the Basic Empathy Scale. *Journal of Adolescence*, *29*(4), 589–611. https://doi.org/10.1016/j.adolescence.2005.08.010

Just, M. A., & Carpenter, P. A. (1999). A Capacity Theory of Comprehension : Indicidual Dirrefences in Working Memory, (January), 1–16.

Kazanas, S. A., & Altarriba, J. (2016). Emotion Word Processing: Effects of Word Type and Valence in Spanish–English Bilinguals. *Journal of Psycholinguistic Research*, *45*(2), 395–406. https://doi.org/10.1007/s10936-015-9357-3

Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual Differences in Language Acquisition and Processing. *Trends in Cognitive Sciences*, *22*(2), 154–169. https://doi.org/10.1016/j.tics.2017.11.006

Kleinginna, P. R., & Kleinginna, A. M. (1981). A categorized list of motivation definitions, with a suggestion for a consensual definition. *Motivation and Emotion*, *5*(3), 263–291. https://doi.org/10.1007/BF00993889

Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The Representation of Abstract Words: Why Emotion Matters. *Journal of Experimental Psychology: General*, *140*(1), 14–34. https://doi.org/10.1037/a0021446

Kousta, S. T., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition*, *112*(3), 473–481. https://doi.org/10.1016/j.cognition.2009.06.007

Kuchinke, L., Võ, M. L. H., Hofmann, M., & Jacobs, A. M. (2007). Pupillary responses during lexical decisions vary with word frequency but not emotional valence. *International Journal of Psychophysiology*, *65*(2), 132–140. https://doi.org/10.1016/j.ijpsycho.2007.04.004

Kunzmann, U., & Grühn, D. (2005). Age differences in emotional reactivity: The sample case of sadness. *Psychology and Aging*, *20*(1), 47–59. https://doi.org/10.1037/0882-7974.20.1.47

Kuperman, V., Estes, Z., Brysbaert, M., & Warriner, A. B. (2014). Emotion and language: Valence and arousal affect word recognition. *Journal of Experimental …*, (4). Retrieved from http://psycnet.apa.org/psycarticles/2014-03909-001

Kuperman, V., & Van Dyke, J. a. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language*, *65*(1), 42–73. https://doi.org/10.1016/j.jml.2011.03.002

Kuppens, P., Tuerlinckx, F., Yik, M., Koval, P., Coosemans, J., Zeng, K. J., & Russell, J. A. (2017). The Relation Between Valence and Arousal in Subjective Experience Varies With Personality and Culture. *Journal of Personality*, *85*(4), 530–542. https://doi.org/10.1111/jopy.12258

Kuznetsova, A., Brockhoff, P., & Christensen, R. (2016). lmerTest: Tests in Linear Mixed Effects Models. *R Package Version*, *3.0.0*, https://cran.r-project.org/package=lmerTest. Retrieved from https://cran.r-project.org/package=lmerTest

Lang, P. J. (1995). The Emotion Probe: Studies of Motivation and Attention. *American Psychologist*. https://doi.org/10.1037/0003-066X.50.5.372

Larsen, R. J., & Ketelaar, T. (1991). Personality and susceptibility to positive and negative emotional states. *Journal of Personality and Social Psychology*, *61*(1), 132–140. https://doi.org/10.1037/0022-3514.61.1.132

Levenson, R. W., Carstensen, L. L., Friesen, W. V., & Ekman, P. (1991). Emotion, physiology, and expression in old age. *Psychology and Aging*, *6*(1), 28–35. https://doi.org/10.1037/0882-7974.6.1.28

Lewinsohn, P. M., Lobitz, W. C., & Wilson, S. (1973). "Sensitivity" of Depressed Individuals to Aversive Stimuli. *Journal of Abnormal Psychology*, *81*(3), 259–263.

Long, L. S., & Titone, D. A. (2007). Psychopathy and verbal emotion processing in non-incarcerated males. *Cognition and Emotion*, *21*(1), 119–145. https://doi.org/10.1080/02699930600551766

Mackay, D. G., Shafto, M., Taylor, J. K., Marian, D. E., Abrams, L., & Dyer, J. R. (2004). Relations between emotion, memory, and attention: Evidence from taboo Stroop, lexical decision, and immediate memory tasks. *Memory & Cognition*, *32*(3), 474–488. https://doi.org/10.3758/BF03195840

Macleod, J. W., Lawrence, M. a, McConnell, M. M., Eskes, G. a, Klein, R. M., & Shore, D. I. (2010). Appraising the ANT: Psychometric and theoretical considerations of the Attention Network Test. *Neuropsychology*, *24*(5), 637–651. https://doi.org/10.1037/a0019803

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research : JSLHR*, *50*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

Mcwhirter, B. T., Besett-Alesch, T. M., Horibata, J., & Gat, I. (2002). Loneliness in High Risk Adolescents: The Role of Coping, Self-Esteem, and Empathy. *Journal of Youth Studies*, *5*(1), 69–84. https://doi.org/10.1080/13676260120111177

Méndez-Bértolo, C., Pozo, M. A., & Hinojosa, J. A. (2011). Word frequency modulates the processing of emotional words: Convergent behavioral and electrophysiological data. *Neuroscience Letters*, *494*(3), 250–254. https://doi.org/10.1016/j.neulet.2011.03.026

Monnier, C., & Syssau, A. (2014). Affective Norms for French Words (FAN). *Behavior Research Methods*, *46*(4), 1128–1137. https://doi.org/10.3758/s13428-013-0431-1

Moors, A., De Houwer, J., Hermans, D., Wanmaker, S., van Schie, K., Van Harmelen, A.-L., … Brysbaert, M. (2013). Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words. *Behavior Research Methods*, *45*(1), 169–177. https://doi.org/10.3758/s13428-012-0243-8

Mor, N., & Winquist, J. (2002). Self-focused attention and negative affect: A meta-analysis. *Psychological Bulletin*, *128*(4), 638–662. https://doi.org/10.1037/0033-2909.128.4.638

O'Connor, L. E., Berry, J. W., Lewis, T., Mulherin, K., & Crisotomo, P. S. (2007). Empathy and depression: the moral system on overdrive. *Empathy in Mental Ilness*. https://doi.org/10.1017/CBO9780511543753.005

Orne, M. T. (1962). On the Social Psychology of Psychological Experiment: With Particular Reference to Demand Characteristics and Their Implications. *American Psychologist*, *17*(11), 776–783. Retrieved from https://journals.scholarsportal.info/details/0003066x/v17i0011/776_otspottdcati.xml

Ostaszewski, K., & Zimmerman, M. A. (2006). The effects of cumulative risks and promotive factors on urban adolescent alcohol and other drug use: A longitudinal study of resiliency. *American Journal of Community Psychology*, *38*(3–4), 237–249. https://doi.org/10.1007/s10464-006-9076-x

Pearson, C., Janz, T., & Ali, J. (2013). *Mental and substance use disorders in Canada*.

Peirce, R., Frone, M., Russell, M., Cooper, M., & Mudar, P. (2000). A Longitudinal Model of Social Contact, Social Support, Depression, and Alcohol Use. *Health Psychol*, *19*(1), 28–38. https://doi.org/10.1037/0278-6133.19.1.28

Pennebaker, J., Booth, R., & Francis, M. (2007). Linguistic inquiry and word count: LIWC. Austin, TX: liwc net.

Perfetti, C. (2007). Reading Ability: Lexical Quality to Comprehension. *Scientific Studies of Reading*, *11*(4), 357–383. https://doi.org/10.1080/10888430701530730

Ponari, M., Rodriguez-Cuadrado, S., & Vinson, D. (2014). Processing advantage for emotional words in bilingual speakers. *Emotion*, *15*(January), 644–652. https://doi.org/10.1037/emo0000061

Ponari, M., Rodriguez-Cuadrado, S., Vinson, D., Fox, N., Costa, A., & Vigliocco, G. (2014). Processing advantage for emotional words in bilingual speakers. *Emotion*, *15*(January), 644–652. https://doi.org/10.1037/emo0000061

Pratto, F., & John, O. P. (1991). Automatic Vigilance: The Attention-Grabbing Power of Negative Social Information. *Journal of Personality and Social Psychology*, *61*(3), 380–391. https://doi.org/10.1037/0022-3514.61.3.380

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *The Behabioral and Brain Sciences*, *4*, 515–526. https://doi.org/10.1016/j.celrep.2011.1011.1001.7.

Puntoni, S., de Langhe, B., & van Osselaer, S. M. J. (2009). Bilingualism and the Emotional Intensity of Advertising Language. *Journal of Consumer Research*, *35*(6), 1012–1025. https://doi.org/10.1086/595022

Pyszczynski, T., Holt, K., & Greenberg, J. (1987). Depression, self-focused attention, and expectancies for positive and negative future life events for self and others. *Journal of Personality and Social Psychology*, *52*, 994–1001. https://doi.org/10.1037/0022-3514.52.5.994

R Core Team, I. (2017). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna (2014). R Foundation for Statistical Computing Authors R Core Team.

Recio, G., Conrad, M., Hansen, L. B., & Jacobs, A. M. (2014). On pleasure and thrill: The interplay between arousal and valence during visual word recognition. *Brain and Language*, *134*, 34–43. https://doi.org/10.1016/j.bandl.2014.03.009

Redondo, J., Fraga, I., Padrón, I., & Comesaña, M. (2007). The Spanish adaptation of ANEW (affective norms for English words). *Behavior Research Methods*, *39*(3), 600–605. https://doi.org/10.3758/BF03193031

Rickard, N. S. (2004). Intense emotional responses to music: a test of the physiological arousal hypothesis. *Psychology of Music*, *32*(4), 371–388. https://doi.org/10.1177/0305735604046096

Rottenberg, J., & Gotlib, I. (2004). Socioemotional Functioning in Depression. In *Mood Disorders: A Handbook of Science and Practice* (pp. 61–77).

Rottenberg, J., Gross, J. J., & Gotlib, I. H. (2005). Emotion context insensitivity in major depressive disorder. *Journal of Abnormal Psychology*, *114*(4), 627–639. https://doi.org/10.1037/0021-843X.114.4.627

Rottenberg, J., & Hindash, A. C. (2015). Emerging evidence for emotion context insensitivity in depression. *Current Opinion in Psychology*, *4*, 1–5. https://doi.org/10.1016/j.copsyc.2014.12.025

Rush, J., Trivedi, M. H., Ibrahim, H. M., Carmody, T. J., Arnow, B., Klein, D. N., … Keller, M. B. (2003). The 16-Item Quick Inventory of Depressive Symptomatology. *Depression*, *54*(5), 573–583. https://doi.org/10.1016/S0006-3223(03)01866-8

Russell, C. J., & Bobko, P. (1992). Moderated Regression Analysis and Likert Scales Too Coarse for Comfort. *Journal of Applied Psychology*, *77*(3), 336–342.

Russell, J. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*. Retrieved from http://psycnet.apa.org/psycinfo/1981-25062-001

Sandell, K., & Bornäs, H. (2017). Functioning Numbness Instead of Feelings as a Direction: Young Adults' Experiences of Antidepressant Use. *Sociology*, *51*(3), 543–558. https://doi.org/10.1177/0038038515591947

Santiago-Rivera, A. L., & Altarriba, J. (2002). The Role of Language in Therapy With the Spanish-English Bilingual Client. *Professional Psychology: Research and Practice*, *33*(1), 30–38. https://doi.org/10.1037/0735-7028.33.1.30

Schutte, N. S., Malouff, J. M., Bobik, C., Coston, T. D., Greeson, C., Jedlicka, C., … Wendorf, G. (2001). Emotional intelligence and interpersonal relations. *Journal of Social Psychology*, *141*(4), 523–536. https://doi.org/10.1080/00224540109600569

Scott, G., O'Donnell, P., Leuthold, H., & Sereno, S. (2009). Early emotion word processing: evidence from event-related potentials. *Biological Psychology*, *80*(1), 95–104. https://doi.org/S0301-0511(08)00073-2 [pii]\r10.1016/j.biopsycho.2008.03.010

Scott, G., O'Donnell, P., & Sereno, S. (2012). Emotion words affect eye fixations during reading. *Journal of Experimental Psychology: Learning Memory and Cognition*, *38*(3), 783–792. https://doi.org/10.1037/a0027209

Segalowitz, N., Trofimovich, P., Gatbonton, E., & Sokolovskaya, A. (2008). Feeling affect in a second language: The role of word recognition automaticity. *The Mental Lexicon*, *3*(1), 47–71. https://doi.org/10.1075/ml.3.1.05seg

Sharot, T. (2011). The optimism bias. *Current Biology*, *21*(23), R941–R945. https://doi.org/10.1016/j.cub.2011.10.030

Sheikh, N. A., & Titone, D. A. (2013). Sensorimotor and linguistic information attenuate emotional word processing benefits: An eye-movement study. *Emotion, 13*(6), 1107–1121. https://doi.org/10.1037/a0032417

Sheikh, N., & Titone, D. (2015). The embodiment of emotional words in a second language: An eye-movement study. *Cognition & Emotion*, (April), 1–13. https://doi.org/10.1080/02699931.2015.1018144

Shields, M., Gorber, S. C., Janssen, I., & Tremblay, M. S. (2011). Bias in self-reported estimates of obesity in Canadian health surveys: An update on correction equations for adults. *Health Reports*, *22*(3).

Sigmon, S. T., & Nelson-Gray, R. O. (1992). Sensitivity to Aversive Events in Depression: Antecedent, Concomitant, or Consequent? *Journal of Psychopathology and Behavioral Assessment*, *14*(3), 225–246.

Skehan, P. (1991). Individual Differences in Second Language Learning. *Studies in Second Language Acquisition*, *13*(02), 275. https://doi.org/10.1017/S0272263100009979

Sloan, D. M., Strauss, M. E., Quirk, S. W., & Sajatovic, M. (1997). Subjective and expressive emotional responses in depression. *Journal of Affective Disorders*, *46*, 135–141. https://doi.org/https://doi.org/10.1016/S0165-0327(97)00097-9

Sloan, D. M., Strauss, M. E., & Wisner, K. L. (2001). Diminished Response to Pleasant Stimuli by Depressed Women. *Journal of Abnormal Psychology*, *110*(3), 488–493. https://doi.org/I0.1037//0021-843X.110.3.488

Smith, D. P., Hillman, C. H., & Duley, A. R. (2005). Influences of Age on Emotional Reactivity During Picture Processing. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *60*(1), P49–P56. https://doi.org/10.1093/geronb/60.1.P49

Soares, A. P., Comesaña, M., Pinheiro, A. P., Simões, A., & Frade, S. (2012). Behavior and Research Methods Corresponding author : *Behavior Research Methods*, *44*(1), 256–269. https://doi.org/10.3758/s13428-011-0131-7

Stadthagen-González, H., Ferré, P., Pérez-Sánchez, M. A., Imbault, C., & Hinojosa, J. A. (2017). Norms for 10,491 Spanish words for five discrete emotions: Happiness, disgust, anger, fear, and sadness. *Behavior Research Methods*. https://doi.org/10.3758/s13428-017-0962-y

Stadthagen-Gonzalez, H., Imbault, C., Pérez Sánchez, M. A., & Brysbaert, M. (2017). Norms of valence and arousal for 14,031 Spanish words. *Behavior Research Methods*, 1–45. https://doi.org/10.3758/s13428-015-0700-2

Statistics Canada. (2013). *National Household Survey*. *Immigration and Ethnocultural Diversity in Canada* (Vol. 99-010-X20). https://doi.org/99-010-X2011001

Sutton, T. M., Altarriba, J., Gianico, J. L., & Basnight-Brown, D. M. (2007). The automatic access of emotion: Emotional Stroop effects in Spanish-English bilingual speakers. *Cognition and Emotion*, *21*(5), 1077–1090. https://doi.org/10.1080/02699930601054133

Thoma, P., Norra, C., Juckel, G., Suchan, B., & Bellebaum, C. (2015). Performance monitoring and empathy during active and observational learning in patients with major depression. *Biological Psychology*, *109*, 222–231. https://doi.org/10.1016/j.biopsycho.2015.06.002

Tsai, J. L., Knutson, B., & Fung, H. H. (2006). Cultural variation in affect valuation. *Journal of Personality and Social Psychology*, *90*(2), 288–307. https://doi.org/10.1037/0022-3514.90.2.288

Velez-Uribe, I., & Rosselli, M. (2019). The auditory and visual appraisal of emotion-related words in Spanish–English bilinguals. *Bilingualism: Language and Cognition*, *22*(1), 30–46. https://doi.org/10.1017/S1366728917000517

Vinson, D., Ponari, M., & Vigliocco, G. (2014). How does emotional content affect lexical processing? *Cognition & Emotion*, *28*(4), 737–746. https://doi.org/10.1080/02699931.2013.851068

Võ, M. L. H., Jacobs, A. M., Kuchinke, L., Hofmann, M., Conrad, M., Schacht, A., & Hutzler, F. (2008). The coupling of emotion and cognition in the eye: Introducing the pupil old/new

effect. *Psychophysiology*, *45*(1), 130–140. https://doi.org/10.1111/j.1469-8986.2007.00606.x

Wang, Y. guang, Wang, Y. qiang, Chen, S. lin, Zhu, C. yan, & Wang, K. (2008). Theory of mind disability in major depression with or without psychotic symptoms: A componential view. *Psychiatry Research*, *161*(2), 153–161. https://doi.org/10.1016/j.psychres.2007.07.018

Waris, O., Soveri, A., Ahti, M., Hoffing, R. C., Ventus, D., Jaeggi, S. M., … Laine, M. (2017). A latent factor analysis of working memory measures using large-scale data. *Frontiers in Psychology*, *8*(JUN), 1–14. https://doi.org/10.3389/fpsyg.2017.01062

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191–1207. https://doi.org/10.3758/s13428-012-0314-x

Warriner, A. B., Shore, D. I., Schmidt, L. A., Imbault, C. L., & Kuperman, V. (2017). Sliding into happiness: A new tool for measuring affective responses to words. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *71*(1), 71–88. https://doi.org/10.1037/cep0000112

Watson, D., Clark, L. A., & Carey, G. (1988). Positive and negative affectivity and their relation to anxiety and depressive disorders. *Journal of Abnormal Psychology*, *97*(3), 346–353. https://doi.org/10.1037/0021-843X.97.3.346

Weisleder, A., & Fernald, A. (2013). Talking to Children Matters: Early Language Experience Strengthens Processing and Builds Vocabulary. *Psychological Science*, *24*(11), 2143–2152. https://doi.org/10.1177/0956797613488145

Westbury, C., Keith, J., Briesemeister, B. B., Hofmann, M. J., & Jacobs, A. M. (2015). Avoid violence, rioting, and outrage; approach celebration, delight, and strength: Using large text corpora to compute valence, arousal, and the basic emotions. *The Quarterly Journal of Experimental Psychology*, *0*(0), 1–24. https://doi.org/10.1080/17470218.2014.970204

Wickham, H., & Chang, W. (2016). Create elegant data visualisations using the grammar of graphics. https://doi.org/10.1093/bioinformatics/btr406

Williams, J. M., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and

psychopathology. *Psychological Bulletin*, *120*(1), 3–24. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8711015

Winskel, H. (2013). The emotional Stroop task and emotionality rating of negative and neutral words in late Thai-English bilinguals. *International Journal of Psychology*. Taylor & Francis. https://doi.org/10.1080/00207594.2013.793800

Wolkenstein, L., Schönenberg, M., Schirm, E., & Hautzinger, M. (2011). I can see what you feel, but i can't deal with it: Impaired theory of mind in depression. *Journal of Affective Disorders*, *132*(1–2), 104–111. https://doi.org/10.1016/j.jad.2011.02.010

Wolter, B. (2001). COMPARING THE L1 AND L2 MENTAL LEXICON A Depth of Individual Word Knowledge Model. *Studies in Second Language Acquisition*, *23*(1), S0272263101001024. https://doi.org/10.1017/S0272263101001024

Zuroff, D., Koestner, R., Moskowitz, D. S., McBride, C., Marshall, M., & Bagby, M. (2007). Autonomous motivation for therapy: A new common factor in brief treatments for depression. *Psychotherapy Research*, *17*(2), 137–148. https://doi.org/10.1080/10503300600919380