



# LUND UNIVERSITY

## Conceptualizing and Measuring Well-Being Using Statistical Semantics and Numerical Rating Scales

Kjell, Oscar

2018

*Document Version:*

Publisher's PDF, also known as Version of record

[Link to publication](#)

*Citation for published version (APA):*

Kjell, O. (2018). *Conceptualizing and Measuring Well-Being Using Statistical Semantics and Numerical Rating Scales*. Lund University.

*Total number of authors:*

1

### General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117  
221 00 Lund  
+46 46-222 00 00





# Conceptualizing and Measuring Well-Being Using Statistical Semantics and Numerical Rating Scales

Oscar N.E. Kjell



**LUND**  
UNIVERSITY

DOCTORAL DISSERTATION

by due permission of the Faculty of Social Sciences, Lund University, Sweden.  
To be defended in Kulturen's auditorium, Lund, the 13<sup>th</sup> of April 2018 at 13.15.

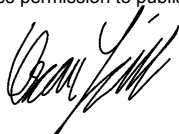
*Faculty opponent*

H. Andrew Schwartz, Stony Brook University

Organization LUND UNIVERSITY	Document name	
	Date of issue	
Author(s): Oscar N.E. Kjell	Sponsoring organization	
Title and subtitle: Conceptualizing and Measuring Well-Being Using Statistical Semantics and Numerical Rating Scales		
<p><b>Abstract</b></p> <p>How to define and measure individuals' well-being is important, as this has an impact on both research and society at large. This thesis concerns how to define and measure the self-reported well-being of individuals, which involves both theorizing as well as developing and applying empirical and statistical methods in order to gain a better understanding of well-being.</p> <p>The first paper critically reviews the literature on well-being. It identifies an individualistic bias in current approaches and accompanying measures related to well-being and happiness; for example, through an over-emphasis on the importance of self-centered aspects of well-being (e.g., the unprecedented focus on satisfaction with life) whilst disregarding the importance of harmony in life, interconnectedness and psychological balance in relation to well-being. It is also discussed how closed-ended well-being measures impose the researchers' values and limit the ability of respondents to express themselves in regard to their perceived well-being.</p> <p>The second paper addresses concerns regarding this individualistic bias by developing the harmony in life scale, which focuses on interconnectedness and psychological balance. In addition, an open-ended approach is developed in the paper, allowing individuals to freely describe their pursuit of well-being by means of open-ended responses analyzed using statistical semantics (including techniques from artificial intelligence such as natural language processing and machine learning). The results show that the harmony in life scale and the traditional satisfaction with life scale form a two-factor model of well-being, where the harmony in life scale explains more unique variance in measures of psychological well-being, stress, depression and anxiety, but not happiness. It is further demonstrated that participants describe their pursuit of harmony in life using words related to interconnectedness (including words such as: peace, balance, cooperation), whereas they describe their pursuit of satisfaction with life using words related to independence (including words such as: money, achievement, fulfillment). It is concluded that the harmony in life scale complements the satisfaction with life scale for a more comprehensive understanding of subjective well-being.</p> <p>The third paper focuses on developing and evaluating a method for measuring and describing psychological constructs using open-ended questions analyzed by means of statistical semantics rather than closed-ended numerical rating scales. This semantic measures approach is tested and compared with traditional rating scales in nine studies, including two different paradigms involving reports regarding objective stimuli (i.e., the evaluation of facial expressions) and reports regarding subjective states (i.e., the self-reporting of harmony in life, satisfaction with life, depression and worry). The results indicate that semantic measures encompass higher, or competitive, levels of reliability and validity compared to traditional numerical rating scales. In addition, semantic measures appear to be better suited for differentiating between psychological constructs, such as harmony in life versus satisfaction with life as well as depression versus worry.</p> <p>In this thesis, the findings from these three papers are elaborated and integrated into two independent perspectives. The first perspective focuses on the theoretical and empirical differences between harmony in life and satisfaction with life within a context of societal and national progress. It is concluded that harmony in life complements satisfaction with life. The second perspective focuses on the open-ended, statistical semantics approach. It is proposed that statistical semantics may beneficially be used more widely as a research tool within psychological research.</p>		
Key words: Well-being, Harmony in life, Satisfaction with life, Statistical semantics, Latent semantic analysis.		
Classification system and/or index terms (if any)		
Supplementary bibliographical information		Language: English
ISSN and key title		ISBN (Print): 978-91-7753-592-8 ISBN (Electronic): 978-91-7753-593-5
Recipient's notes	Number of pages	Price
	Security classification	

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature



Date 2018-02-26

# Conceptualizing and Measuring Well-Being Using Statistical Semantics and Numerical Rating Scales

Oscar N.E. Kjell



**LUND**  
UNIVERSITY

Coverphoto by Oscar Kjell

Copyright Paper I American Psychological Association

Copyright Paper II Springer

Copyright Paper II and Thesis Oscar N.E. Kjell


Faculty of Social Sciences  
Department of Psychology

ISBN (Print): 978-91-7753-592-8

ISBN (Electronic): 978-91-7753-593-5

Printed in Sweden by Media-Tryck, Lund University  
Lund 2018



MADE IN SWEDEN 

Media-Tryck is an environmental-  
ly certified and ISO 14001 certified  
provider of printed material.  
Read more about our environmental  
work at [www.mediatryck.lu.se](http://www.mediatryck.lu.se)





# Table of Contents

Acknowledgement.....	8
Abstract .....	10
Sammanfattning .....	12
List of Papers.....	15
List of Abbreviations.....	16
<b>Introduction .....</b>	<b>17</b>
Overview of the Papers .....	18
Format and Aims of the Thesis .....	21
References .....	23
<b>Perspective I</b>	
Harmony in Life Complements Satisfaction with Life in Measuring Subjective Well-Being for National Progress* .....	25
Abstract .....	25
National Progress Indicators and Well-Being.....	27
HIL Meaningfully Complements SWL.....	29
A comprehensive and representative view of well-being.....	29
Nurturing selflessness rather than competing self-interests .....	30
Addressing the importance of interconnectedness .....	31
HIL and SWL as Indicators of National Progress.....	32
Well-being and economics .....	32
Well-being and society .....	33
Well-being and the environment .....	35
The sustainability of well-being.....	36
Concluding Remarks.....	37
References.....	38
<b>Perspective II</b>	
Statistical Semantics: Measuring and Describing Psychological Phenomena through Natural Language* .....	43
Abstract .....	43
Types of Automated Text Analyses .....	45

Word frequency strategies.....	45
Statistical semantics: Word pattern analyses.....	46
Evaluation of word count strategies and statistical semantics.....	48
Using Semantic Representations in Analyses .....	50
Semantic spaces and semantic representations.....	50
Semantic similarity.....	51
Semantic-numeric correlations and semantic predictions using training.....	51
Describing the meaning of words through visualization .....	52
Applying Statistical Semantics in Psychological Research .....	54
Individuals' accounts related to psychological phenomena .....	54
Psychological aspects of naturally occurring (big) data .....	57
Enhancing experimental control and effects .....	59
The Potential of Statistical Semantics.....	60
Objective, systematic and quantitative qualities.....	60
Validity and reliability.....	61
Flexibility in collecting and analyzing data.....	61
Replications and advancements.....	62
Limitations and challenges .....	63
Conclusion.....	64
References.....	64

### **Paper I**

Sustainable Well-Being: A Potential Synergy Between Sustainability and Well-Being Research

### **Paper II**

The Harmony in Life Scale Complements the Satisfaction with Life Scale: Expanding the Conceptualization of the Cognitive Component of Subjective Well-Being

### **Paper III**

Semantic Measures: Using Natural Language Processing to Measure, Differentiate and Describe Psychological Constructs

### **Appendix A**

Candidate Item Pool for the Harmony in Life Scale

### **Appendix B**

Semantic Excel: An Introduction to a User-Friendly Online Software Application for Statistical Analyses of Text Data

# Acknowledgement

I am happy to have been given the opportunity to write a PhD thesis, a time filled with stimulating work and the privileged support of others. The ideas and work presented in this thesis would not have been possible had it not been for the support of others. For all of this, I feel very grateful!

I am very fortunate to have had dedicated and knowledgeable supervisors. First of all, I would like to show my gratitude to Sverker Sikström, my main supervisor. From the very start, Sverker showed an interest in my ideas about well-being, but also enthusiastically shared his own research experiences. Your creativity in tackling research with new and unconventional ideas has been very educational for me. I appreciated your course in statistical semantics and the patience you displayed when teaching me these analytic techniques. I want to thank Daiva Daukantaitė for being an excellent supervisor. You have encouraged me to believe in my ideas and research, and for that I am very grateful. I am thankful for all the statistical advice and insightful feedback you have provided and for the course you gave in positive psychology. Sverker and Daiva, your dedication, professionalism and friendliness really made work incredibly fun and inspiring. I also want to thank Kate Heffernon, my third supervisor. You gave me valuable comments when developing the harmony in life scale. Being located in London, she was mostly involved in the beginning of my PhD studies as well as the time leading up to them. I am also thankful for your teaching during my MSc studies, which inspired me to want to continue in academia.

I thankfully acknowledge the financial support for my research from the Swedish Research Council and the Helge Ax:son Johnson Foundation. There are many colleagues at the Department of Psychology at Lund University who have inspired me in different ways. Specifically, I want to thank: Martin Bäckström for providing insightful feedback at my half-time review seminar and final review seminar, as well as arranging valuable statistics courses and workshops at the department. Fredrik Björklund for the valuable feedback he provided at my half-time review seminar. I would also like to thank my colleagues and friends: Per Davidson, Arvid Erlandsson, Holly Knapton, David Marcusson-Clavertz (especially for his comments on perspective II), Henrik Levinsson, Martin Hassler Hallstedt and Jean-Christophe Rohner for advice and stimulating discussions about research at work and/or at the pub.

I want to show my gratitude to colleagues outside of Lund University. I would like to thank: Danilo Garcia for interesting collaborations, Itai Ivztan and Tim Lomas for research opportunities as well as involving me in teaching at the MSc in Applied Positive Psychology at the University of East London. Sam Thompson for his excellent supervision during my MSc dissertation and during the process of getting

it published. Deirdre Walsh, Karen Beckwith Wadey and Vera Hegarty for giving feedback on my earlier work. I also want to thank Angela Clow and Tobias Stalder for making me interested in research during my BSc in Psychology at the University of Westminster and for Tobias' friendship and continued advice on academic matters.

My deepest gratitude goes to my family and friends for their loving support; without you the thesis would not have been possible or enjoyable. I would especially like to thank: my parents Anita and Per-Inge Kjell who have always inspired me to pursue work I find interesting, enjoyable and valuable; my eldest sister (and nutritional health coach) Victoria Carinci and her family Daniele, Julia and Ceasar and my sister Johanna Kjell; they have all contributed to my well-being and inspiration. My future mother-in-law Karin Lundberg for helping us in everyday life with practical things like babysitting and encouragement during times of a high work load. Fernando Arias for supporting and encouraging me in so many ways; we had loads of fun behind the school desk from a young age and he gave me valuable feedback and support during the writing process of my first paper on sustainable well-being. David Peković for being the reason I started studying in London and for making that time exciting and memorable. Ola Aronsson, Henrik Eriksson, Karin Krifors, Karl Kronkvist, Anton Pöllänen and Richard Scott, for asking thoughtful questions and showing an interest in my work. But above all, I am grateful that you have all given me numerous laughs and adventures outside of academia.

The ones most important for my academic work as well as in my life are my wife-to-be Katarina and our son Olle. Your love and support have given me profound harmony in life.

## Abstract

How to define and measure individuals' well-being is important, as this has an impact on both research and society at large. This thesis concerns how to define and measure the self-reported well-being of individuals, which involves both theorizing as well as developing and applying empirical and statistical methods in order to gain a better understanding of well-being.

The first paper critically reviews the literature on well-being. It identifies an individualistic bias in current approaches and accompanying measures related to well-being and happiness; for example, through an over-emphasis on the importance of self-centered aspects of well-being (e.g., the unprecedented focus on satisfaction with life) whilst disregarding the importance of harmony in life, interconnectedness and psychological balance in relation to well-being. It is also discussed how closed-ended well-being measures impose the researchers' values and limit the ability of respondents to express themselves in regard to their perceived well-being.

The second paper addresses concerns regarding this individualistic bias by developing the harmony in life scale, which focuses on interconnectedness and psychological balance. In addition, an open-ended approach is developed in the paper, allowing individuals to freely describe their pursuit of well-being by means of open-ended responses analyzed using statistical semantics (including techniques from artificial intelligence such as natural language processing and machine learning). The results show that the harmony in life scale and the traditional satisfaction with life scale form a two-factor model of well-being, where the harmony in life scale explains more unique variance in measures of psychological well-being, stress, depression and anxiety, but not happiness. It is further demonstrated that participants describe their pursuit of harmony in life using words related to interconnectedness (including words such as: *peace*, *balance*, *cooperation*), whereas they describe their pursuit of satisfaction with life using words related to independence (including words such as: *achievement*, *money*, *fulfillment*). It is concluded that the harmony in life scale complements the satisfaction with life scale for a more comprehensive understanding of subjective well-being.

The third paper focuses on developing and evaluating a method for measuring and describing psychological constructs using open-ended questions analyzed by means of statistical semantics rather than closed-ended numerical rating scales. This *semantic measures approach* is tested and compared with traditional rating scales in nine studies, including two different paradigms involving reports regarding objective stimuli (i.e., the evaluation of facial expressions) and reports regarding subjective states (i.e., the self-reporting of harmony in life, satisfaction with life, depression and worry). The results indicate that semantic measures encompass

higher, or competitive, levels of reliability and validity compared to traditional numerical rating scales. In addition, semantic measures appear to be better suited for differentiating between psychological constructs, such as harmony in life versus satisfaction with life as well as depression versus worry.

In this thesis, the findings from these three papers are elaborated and integrated into two independent perspectives. The first perspective focuses on the theoretical and empirical differences between harmony in life and satisfaction with life within a context of societal and national progress. It is concluded that harmony in life complements satisfaction with life. The second perspective focuses on the open-ended, statistical semantics approach. It is proposed that statistical semantics may beneficially be used more widely as a research tool within psychological research.

## Sammanfattning

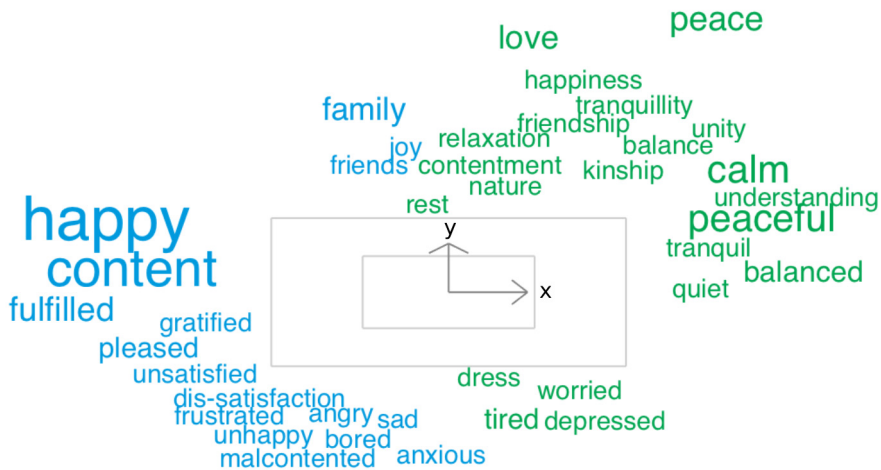
Hur man beskriver och eftersträvar välbefinnande är viktigt, eftersom det påverkar både forskning och samhället i stort. Denna avhandling handlar om hur välbefinnande definieras och mäts, vilket både innefattar en kritisk diskussion kring befintliga teorier om välbefinnande samt en utveckling av empiriska metoder och statistiska verktyg för att öka vår förståelse gällande hur personer uppfattar och eftersträvar välbefinnande.

Den första artikeln granskar välbefinnandelitteraturen. Den diskuterar hur individens sökande efter välbefinnande kan kombineras med ett hållbart leverne, vilket handlar om hur individen uppfyller sina behov utan att äventyra kommande generationers förutsättningar att uppfylla sina. Rådande teorier om välbefinnande och lycka har ett individualistiskt fokus, vilket även återspeglas i de mätinstrument som används idag. Litteraturen tenderar att ensidigt fokusera på betydelsen av det egna jaget, samtidigt som vikten av samhörighet och ömsesidigt beroende till andra människor och naturen blir sekundära. Ett tydligt exempel på detta är välbefinnandebegreppet Livstillfredsställelse, vilket oftast mäts genom att individen får svara på frågor om i vilken uträkning omgivningen stämmer överens med dennes ideala förväntningar. Omgivningen ska alltså ensidigt formas, så långt det bara går, till att passa individens önsknings. Detta fokus bortser från vikten av balans och ömsesidigt beroende i förhållandet mellan individen och omgivningen. Välbefinnande kan alltså även nås genom harmoni och i balans med sin omgivning. Artikeln diskuterar även hur slutna svarsalternativ i de numeriska skattningsskalor som vanligtvis används för att mäta välbefinnande begränsar svarandens möjlighet att uttrycka sig och därmed även den information som kommer forskningen till del.

Den andra artikeln behandlar välbefinnandeforskningens individualistiska fokus genom att utveckla mätinstrumentet *Harmoni i livet*. Detta instrument syftar till att komplettera det ensidiga fokuset på jaget i de mätinstrumentet som idag vanligtvis används för att mäta välbefinnande. Harmoni i livet fokuserar på ömsesidigt beroende och psykologisk balans. Vi utvecklar även en metod där personer kan beskriva sin strävan efter välbefinnande med ord som analyseras kvantitativt med hjälp av statistisk semantik och metoder från artificiell intelligens. Resultaten visar att harmoni i livet och tillfredsställelse med livet utgör två unika komponenter av välbefinnande, där harmoni i livet har en starkare koppling till psykologiskt välbefinnande, stress, depression och ångest, men inte lycka. Även resultaten från de beskrivande orden uppvisar en tydlig skillnad i hur deltagare beskriver harmoni och tillfredsställelse (se Figur 1 nedan). Personer beskriver sin strävan efter harmoni med ord som knyter an till samhörighet och balans (till exempel *sammanhållning*, *balans*, *samarbete*, *tillsammans* och *förståelse*), medan de beskriver sin strävan efter tillfredsställelse med ord som fokuserar på det egna jaget (till exempel *prestation*,







Tillfredsställelse-ord vs. Harmoni-ord

**Figur 2.**

X-axeln visar statistiskt signifikanta engelska ord som skiljer mellan hur individer beskriver sin upplevelse av tillfredsställelse med livet (blåa ord) och harmoni i livet (gröna ord). På y-axeln plottas orden enligt hur semantiskt nära de är begreppet harmoni in livet. Mer frekventa ord har större teckensnitt. Bilden beskrivs i detalj i avhandlingens tredje manus.

Avhandlingens introduktion diskuterar och utvecklar resultaten från de tre artiklarna i två oberoende perspektiv. Det första perspektivet lägger fokus på de teoretiska och empiriska skillnaderna mellan harmoni i livet och livstillfredsställelse. Dessa diskuteras inom en samhällelig kontext och i relation till framgångs- och välståndsex. Sökandet efter harmoni i livet och livstillfredsställelse relateras till ekonomiska, sociala, miljö- och hållbarhetsaspekter. Här diskuteras även vikten av att komplettera det rådande fokuset på jaget och egen livstillfredsställelse i strävan efter välbefinnande.

Det andra perspektivet lägger fokus på den statistiska semantikmetoden och utvecklingen av semantiska mätinstrument. Statistisk semantik och dess tillämpningsområden inom psykologisk forskning beskrivs. Potentialen för statistisk semantik som ett verktyg inom psykologisk forskning diskuteras.

## List of Papers

- Paper I Kjell, O. N. E. (2011). Sustainable Well-Being: A Potential Synergy Between Sustainability and Well-Being Research. *Review of General Psychology*, 15(3), 255-266. doi:10.1037/a0024603  
(Published paper reprinted with permission from American Psychological Association)
- Paper II Kjell, O. N. E., Daukantaitė, D., Hefferon, K., & Sikström, S. (2016). The Harmony in Life Scale Complements the Satisfaction with Life Scale: Expanding the Conceptualization of the Cognitive Component of Subjective Well-Being. *Social Indicators Research*, 126(2), 893-919. doi:10.1007/s11205-015-0903-z  
(Published paper reprinted with permission from Springer)
- Paper III Kjell, O.N.E., Kjell, K., Garcia, D., & Sikström, S. (in revision at *Psychological Methods*). Statistically Measuring and Describing Psychological Constructs with Words, not Numbers.

## List of Abbreviations

HIL	Harmony in life
HILS	Harmony in life scale
SWL	Satisfaction with life
SWLS	Satisfaction with life scale
SWB	Subjective well-being
PWB	Psychological well-being
LSA	Latent semantic analysis
LDA	Latent Dirichlet allocation
DLA	Differential language analysis
LIWC	Linguistic inquiry and word count

# Introduction

The way we think about, or define, well-being influences the way we pursue it. If we focus on a form of well-being defined as satisfaction with life, we are likely to pursue achievements, goals, work, career, money and pleasures (Kjell, Daukantaitė, Hefferon, & Sikström, 2016). If, on the other hand, we focus on harmony in life, we are likely to pursue peace, balance, agreement, unity, friendship and cooperation (Kjell et al., 2016). The way we define well-being also influences research and society. This influence may play out something like this: The way in which well-being is defined has a direct impact on the construction of questions for measuring it. If we define well-being as satisfaction with life, then the questions composing well-being questionnaires will focus on satisfaction and achievements rather than, for example, harmony and balance. As a result, well-being questionnaires influence research by being used for evaluating the effectiveness of therapies or interventions aimed at increasing well-being. It is thus possible that an intervention primarily increasing an individual's sense of peace and balance, but not his or her sense of achievement and satisfaction, is deemed irrelevant if it has been evaluated based on satisfaction with life rather than harmony in life. Eventually, these research findings will impact the society at large in the form of guidance for social policies, therapies, teaching, parenting and so on.

How to define and measure well-being is the main topic of this thesis. The term *well-being* is usually used as an umbrella term broadly referring to “optimal psychological functioning and experience” (Ryan & Deci, 2001, p. 142). As such, well-being is complex. Recently, Linton, Dieppe, and Medina-Lara (2016) reviewed 99 well-being measures and concluded that there are many approaches for how to define and measure well-being, with even more dimensions categorized to fall within the realm of well-being. One of the most common approaches, referred to as *subjective well-being* (Diener, 1984), emphasizes the subjective experience and understanding of well-being. The aim is to enable each individual to decide for her- or himself what well-being entails. In contrast, *psychological well-being*, another

common approach,<sup>1</sup> offers a stricter definition of well-being for individuals by means of six predefined dimensions (Ryff, 1989).

Linton et al. (2016) found 196 different dimensions or underlying aspects of well-being that were measured across the well-being measures. The six dimensions in the approach concerning psychological well-being include: i) self-acceptance, ii) positive relationships, iii) autonomy, iv) environmental mastery, v) purpose in life, and vi) personal growth (Ryff, 1989). The subjective well-being approach focuses on an affect component including positive and negative affect, as well as a cognitive component, which concerns how individuals evaluate their life (Diener, 1984). The most common dimension for the cognitive component is satisfaction with life (Diener, Emmons, Larsen, & Griffin, 1985). However, this thesis shows how harmony in life may complement satisfaction with life (Kjell, 2011; Kjell et al., 2016).

The thesis includes considerations concerning theoretical approaches, aspects related to defining, conceptualizing and pursuing well-being as well as methodological and statistical aspects related to how it should be measured. The statistical methods not only involve traditional frequentist statistics, but also techniques from artificial intelligence, including natural language processing and machine learning. The three papers composing the thesis are described below, followed by a description concerning the format of the two perspectives used for integrating and discussing the included papers and related research.

## Overview of the Papers

The first paper reviews common definitions and associated measures related to well-being (Kjell, 2011). The paper identifies an individualistic, self-involved, self-serving bias in current definitions and associated measures related to well-being. For example, one of the most commonly used well-being measures, the satisfaction with life scale (Diener et al., 1985), is based on a definition of well-being focusing on a state where an individual's surroundings and circumstances correspond to his or her ideal expectations, which is held to be the highest, all-encompassing form of happiness and well-being. The respondents are encouraged to only consider themselves and their own personal wishes and desires without considering other people or the natural environment. That is, one person's satisfaction may easily be another person's dissatisfaction, or, for that matter, using up scarce natural resources

---

<sup>1</sup> Linton, Dieppe and Medina-Lara (2016) found that in the literature, the *subjective well-being* approach is one of the two most common approaches and that the *psychological well-being* approach is one of seven well-being approaches commonly referred to in the literature.

(Kjell, 2011). To balance this view, this paper expresses the need for a complementary approach incorporating a more interdependent and contextualized view of well-being – a view that accounts for contextual factors, such as the importance of other people and nature. In this regard, a synergy between research on well-being and on sustainability is proposed, highlighting prospects of combining the pursuit of well-being and sustainable living. One of the practical suggestions is to complement the unprecedented focus on satisfaction with life with a focus on harmony in life and psychological balance.

The paper also raises methodological concerns regarding the way in which well-being is measured. It is traditionally held that using the satisfaction with life scale does not force respondents to think about well-being in any specific way, but instead allows them to decide for themselves which aspects of life they deem important. However, in the paper it is pointed out that the focus on satisfaction with life is in and of itself imposed. Respondents need to adhere to the focus on satisfaction with life and the items comprising the satisfaction with life scale. In other words, they are (at least implicitly) required to think about well-being in terms of what they have, compared to what they want (i.e., their ideal expectations). Another problem with this approach is that we know very little about what individuals think about when answering the items. That is, the answers to these broad, unspecific items (for example, *I am satisfied with my life*) are provided using numbers from the numerical response scale ranging from 1 = *strongly disagree* to 7 = *strongly agree*. This means that these responses do not indicate whether the respondent was considering his or her family, house, work, all three or something entirely different. Hence, it is difficult to understand what respondents were thinking about when answering these closed-ended well-being scales.

The second paper develops the *harmony in life scale* (Kjell et al., 2016), which is a well-being measure focusing on psychological balance and interconnectedness (Appendix A displays all items developed as potential items for the final harmony in life scale). From a perspective of well-being, this paper addresses critique identified in the first paper related to the self-centered view of well-being. It demonstrates that harmony in life and satisfaction with life form a two-factor model representing distinct aspects of well-being, even though they are strongly correlated. For example, when comparing harmony in life and satisfaction with life, the former is more related to psychological well-being and the latter is more related to happiness. Further, compared to satisfaction with life, harmony in life explains a more unique variance in depression, anxiety and stress.

To empirically examine which aspects of life individuals consider in relation to pursuing harmony in life versus satisfaction with life, we employed a method that allowed for open-ended word answers, which were then analyzed using techniques from artificial intelligence, including natural language processing and machine

learning. In particular, we used statistical semantics based on Latent Semantic Analysis, which is a method where the words in a language are described using numerical values in a high-dimensional space. These numerical values enable statistical analyses of word responses. So, in our studies, individuals were asked to write down words they associated with the pursuit of harmony, satisfaction, psychological well-being and happiness. Statistical analyses of these words revealed that words generated in relation to the four different well-being constructs differed significantly. Words describing the pursuit of harmony were semantically similar to words describing the pursuit of psychological well-being, whereas words describing the pursuit of satisfaction were semantically similar to words describing the pursuit of happiness. By statistically comparing words generated in relation to harmony versus satisfaction, we furthermore found that individuals relate their pursuit of harmony using words such as *peace* and *balance*, whereas they relate their pursuit of satisfaction using words such as *job* and *money*. Hence, this addresses the methodological concerns regarding which aspects respondents think about in relation to the different well-being constructs.

In the third paper, we develop and evaluate a new way of measuring psychological constructs using open-ended word responses rather than numerical rating scales with closed-ended responses (Kjell, Kjell, Garcia, & Sikström, in revision). Rating scales are made up of items such as *I am in harmony*, which are coupled with closed-ended responses such as *1 = strongly disagree* to *7 = strongly agree* (Kjell et al., 2016). In contrast, we constructed questions allowing for open-ended word responses, which we refer to as semantic questions (e.g., “*Overall in your life, are you in harmony or not?*”). These semantic questions (concerning harmony, satisfaction, depression and worry) enabled respondents to answer them using descriptive words or texts. This is important as people naturally describe complex psychological experiences using open-ended word responses (such as *joyful* or *peaceful*) rather than using closed-ended, forced-choice, numerically based response formats (such as “*7 = strongly agree*”). Using statistical semantics, we found that these word responses may be trained to predict numerical rating scale scores with a high level of accuracy. For example, the words describing an individual’s harmony in life may predict his or her harmony in life scale score remarkably well ( $r = .72, p < .001$ ). This demonstrates that the semantic responses are valid in relation to the numerical rating scales.

To measure the constructs independently from numerical rating scales, we created word norms by asking individuals to describe the to-be-measured constructs (i.e., harmony in life, satisfaction with life, depression and worry). These word norms are used for measuring the semantic similarity to word responses from individual answers to the semantic questions. So, for example, the harmony in life word norm is typically used for measuring the semantic similarity to an individual’s response to a semantic question about harmony in life, and the higher the semantic similarity

between the two sets of words, the higher the degree of harmony in life reported by this individual. Using this method, we demonstrated that semantic measures may be used for *measuring* psychological constructs independent of closed-ended numerical rating scales. The analyses revealed that compared to satisfaction with life, harmony in life exhibits a stronger negative correlation to both depression and worry.

The word responses from semantic questions may also be used for *describing* the constructs. On the x-axes, we may for instance plot words that are significantly different between the answers for two different constructs, while on the y-axes plot words relating to low and high semantic similarity or scores on a numerical rating scale. Plotting words this way suggests that semantic measures (i.e., measures derived from the responses from semantic questions) are better at discriminating between psychological constructs compared to numerical rating scales. In other words, plotting words according to the harmony in life semantic similarity scores tends to highlight harmony words more than satisfaction words, and vice versa for the satisfaction with life semantic similarity scores, whereas the satisfaction with life rating scale scores and the harmony in life rating scale scores do not exhibit this distinct pattern. We found that participants described their experience of harmony in life using words such as *peace*, *balance* and *unity*, whereas they described their high satisfaction with life using words such as *content*, *happy* and *fulfilled*.

It is worth pointing out that this paper addresses several *methodological* problems relevant for psychological science, especially considering the fact that self-reporting of psychological constructs is widely assessed using closed-ended numerical rating scales. Semantic questions enable individuals to freely express themselves when self-reporting their experiences in relation to psychological constructs. In addition, whereas numerical rating scales force individuals to answer on one-dimensional closed-ended numerical rating scales, semantic questions enable multi-descriptive responses. Hence, semantic questions are capable of both *measuring* and *describing* a psychological construct. Overall, the results suggest that semantic measures have a higher, or competitive, level of validity and reliability as compared with numerical rating scales.

## Format and Aims of the Thesis

The papers composing this thesis signify two different insights, each deserving attention in its own right. The first pertains to well-being, including differences between harmony in life and satisfaction with life. The second concerns methodological and statistical aspects of conceptualizing and measuring psychological constructs using words rather than numerical rating scales as the



response format. In order to clearly describe, elaborate and put these two insights in their respective research contexts, as well as to accommodate for the prospect of having two different target groups (i.e., one primarily interested in well-being, the other primarily interested in methodology and statistical semantics), the remainder of the general introduction is written in the form of two perspectives (rather than the more traditional format beginning with a summary of the field followed by a short summary of the papers included).

Perspectives normally consist of a concise account of an important topic, insight or approach that encourages authors to be more thought-provoking compared to when writing more typical articles. In relation to my research, the aim is to summarize and integrate the most pertinent findings in two cohesive narratives that are accessible to an audience with little background knowledge regarding the specific topics. The aspiration is to present my perspective on how the thesis is relevant for psychological science and society at large. I focus in particular on highlighting the emerging conclusions enabled by considering the papers as one entity and in relation to existing research. Hence, each of the perspectives independently summarizes and integrates findings from all three papers and links these findings to related research within a larger context. The two perspectives may also be read separately from one another as well as independently from the papers.

The first perspective focuses on well-being and is entitled *Harmony in Life Complements Satisfaction with Life in Measuring Subjective Well-Being for National Progress*. As the first paper includes a review of the literature on well-being, this perspective focuses more on the differences between harmony in life and satisfaction with life. Further, since I aspire to be concrete and avoid unnecessary abstraction, these differences are discussed within a societal context. An overall aim for this perspective involves demonstrating the importance of carrying out more empirical research on harmony in life.

The second perspective focuses on the statistical and methodological aspects of the papers. It is called *Statistical Semantics: Measuring and Describing Psychological Phenomena through Natural Language*. This perspective describes and evaluates different approaches to automated text analyses. For a thorough description of statistical semantics see Appendix B (entitled *Semantic Excel: An Introduction to a User-Friendly Online Software Application for Statistical Analyses of Text Data*). As this perspective is oriented toward statistics and methodology, it encompasses a broader scope than merely discussing well-being constructs. In this perspective, it is proposed that statistical semantics as a research tool may be used more broadly in psychological science than what has been the case so far.

## References

- Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, 95(3), 542-575. doi:10.1037//0033-2909.95.3.542
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, 49(1), 71-75.
- Kjell, O. N. E. (2011). Sustainable Well-Being: A Potential Synergy Between Sustainability and Well-Being Research. *Review of General Psychology*, 15(3), 255-266. doi:10.1037/a0024603
- Kjell, O. N. E., Daukantaitė, D., Hefferon, K., & Sikström, S. (2016). The Harmony in Life Scale Complements the Satisfaction with Life Scale: Expanding the Conceptualization of the Cognitive Component of Subjective Well-Being. *Social Indicators Research*, 126(2), 893-919. doi:10.1007/s11205-015-0903-z
- Kjell, O. N. E., Kjell, K., Garcia, D., & Sikström, S. (in revision at *Psychological Methods*). Semantic Measures: Using Natural Language Processing to Measure, Differentiate and Describe Psychological Constructs..
- Linton, M.-J., Dieppe, P., & Medina-Lara, A. (2016). Review of 99 self-report measures for assessing well-being in adults: exploring dimensions of well-being and developments over time. *BMJ Open*, 6(7). doi:10.1136/bmjopen-2015-010641
- Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology*, 52, 141-166. doi:10.1146/annurev.psych.52.1.141
- Ryff, C. D. (1989). Happiness is everything, or is it - explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, 57(6), 1069-1081.



# Perspective I

## Harmony in Life Complements Satisfaction with Life in Measuring Subjective Well-Being for National Progress\*

### Abstract

Measures of individuals' self-reported subjective well-being are currently being more used to complement existing national progress indices. Consequently, the way subjective well-being is defined and measured impacts how our society is evaluated and influences the goals we strive for. Whereas subjective well-being is most frequently measured as satisfaction with life, we argue that this is a one-sided focus that warrants attention. A focus on satisfaction with life may lead to an overly narrow emphasis on individualistic self-interest, self-serving needs and independence, whilst not taking the importance of interconnectedness into account. Based on empirical findings, we propose that harmony in life serves as an important complement to the current one-sided focus on satisfaction with life. Harmony in life involves interconnectedness and psychological balance. We discuss differences between satisfaction with life and harmony in life in relation to national progress on the basis of several related perspectives, including economics, society, environment and sustainability. Individuals' evaluations of harmony in life appear to involve them beneficially considering contextual aspects linked to society, sustainability and the environment. We encourage advancing the empirical research on harmony in life since it meaningfully complements satisfaction with life, and strengthens many diverse goals considered important for national progress.

*Keywords:* Harmony in life; Satisfaction with life; Subjective well-being; Happiness; National progress indicators

---

\* This perspective will be submitted for potential publication together with Daiva Daukantaitė and Sverker Sikström.

Asking individuals to report their subjective well-being (SWB) as a way of evaluating national progress has gained in interest among researchers and policymakers (e.g., see Dolan & White, 2007; Stiglitz, Sen, & Fitoussi, 2010). Diener (2000) proposes that SWB measures may complement traditional national progress indicators in order to, for example, better inform public policies. Traditional national progress indicators strongly emphasize economic activity, which at best only serves as an indirect measure of individual well-being. This means that SWB assessments are useful as they capture individuals' own experiences and evaluations in response to different policies (Diener, 2000). Diener, Lucas, Schimmack, and Helliwell (2009) believe that future standard practices in most countries are likely to capture SWB. The fact that more than 40 countries already employ SWB measures in some form (Diener, Oishi, & Lucas, 2015) may be seen as a testament to the current interest in drawing conclusions regarding national progress from the self-reported well-being of individuals. In light of these advances, the way in which SWB is defined and measured will have wide-ranging implications for national progress.

Exactly how we define well-being influences our research and societies at large. The definition governs us in the way we phrase the questions included in well-being measures. In turn, the definition and the measures have an impact on research; for example, in how researchers design and evaluate clinical therapies or positive psychology interventions. Subsequently, these findings guide governmental practices, social policies, therapies, coaching, teaching, parenting and so on. The definition and measures also have an impact on societies as SWB measures are used for evaluating national progress. Diener and Seligman (2004) point out that what is measured by societies is likely to receive more attention and in turn be the very thing to which meaning is attached and ultimately pursued. This raises the question: What exactly should be defined and measured as SWB?

The SWB approach emphasizes the importance of enabling people to decide for themselves what well-being entails (Diener, Sapyta, & Suh, 1998). The aim is thus to avoid forcing a particular view of well-being upon individuals. SWB is conceptualized to encompass an affective component comprising positive and negative affect and a cognitive component comprising life evaluations (Diener, 1984). The cognitive component is *typically* defined as satisfaction with life (SWL), where Diener, Emmons, Larsen, and Griffin (1985) argue that SWL is *the* overarching cognitive construct reflecting happiness and well-being. Further, reports recommending a focus on SWB accounts in the context of national progress one-sidedly tend to promote an SWL focus (e.g., see Diener, 2000; Diener, Inglehart, & Tay, 2013; Diener et al., 2009; Diener et al., 2015; Stiglitz et al., 2010). In contrast, we argue that the current unprecedented focus on SWL within SWB is in itself forcing a particular view of well-being upon individuals and that it represents an overly narrow focus on individualistic self-interests, whilst failing to

adequately capture the interconnectedness of well-being. Based on a body of recent empirical research (e.g., see Delle Fave, Brdar, Freire, Vella-Brodrick, & Wissing, 2011; Delle Fave et al., 2016; Kjell, Daukantaitė, Hefferon, & Sikström, 2016; Kjell, Kjell, Garcia, & Sikström, in revision), we propose that harmony in life (HIL) meaningfully complements SWL by capturing aspects central to the well-being of individuals, including psychological balance and interconnectedness. It is further proposed that HIL is particularly vital for strengthening many of the diverse goals conceptualized as national progress.

## National Progress Indicators and Well-Being

Standard national economic indicators, such as Gross Domestic Product (GDP) and Gross National Product (GNP), measure economic activity. However, they are not constructed to, nor do they, assess the social or economic welfare of countries. For example, Diener and Seligman (2004) point out that decreasing mental health may increase GDP if more money is spent on care. “Paradoxically,” they write, “a mounting problem in well-being might increase economic indicators, and the increase in GDP does not indicate whether the money is spent effectively” (p. 17). Likewise, whereas oil spills may increase GDP by requiring costs related the clean-up, they reduce overall well-being in a country (Costanza et al., 2004). Despite the shortcomings of these economic measures, they have to a great extent misguidedly been used as measures of social and economic welfare (Costanza, Hart, Talberth, & Posner, 2009; Goossens & Mäkipää, 2007; Stiglitz et al., 2010).

To complement mere economic measures, innovative indicators now evaluate progress in terms of sustainability, societal and environmental advances. For example, the Genuine Progress Index captures sustainable economic welfare as opposed to mere economic activity, as it adjusts the GDP by distinguishing between economic activities that diminish versus enhance natural and social capital (Talberth, Cobb, & Slattery, 2007). Another example is the Human Development Index, which is a composite index encompassing three dimensions, including knowledge and education, a long and healthy life and a good standard of living (UNDP, 2015). These indicators certainly unveil a different, more comprehensive, picture of progress compared to simply looking at the GDP. Notably though, current indicators typically comprise objective circumstances (e.g., health, education etc.), but do not incorporate accounts of how individuals subjectively evaluate their life (however, see the Happy Planet Index; Abdallah, Thompson, Michaelson, Marks, & Steuer, 2009).

Diener and Suh (1997) point out that national SWB accounts are valuable as they do not simply reflect *objective circumstances* (e.g., number of teachers, doctors,

nurses or police officers per capita), but capture how individuals evaluate these circumstances. Further, Diener et al. (2009) argue that the fact that SWB reflects how individuals overall evaluate their lives means that these measures should primarily change when there is change in areas that are decisive for individuals. Therefore, they argue that SWB measures possess the potential of offering “information about the relative importance of the various domains in people’s lives—information that is crucial for making decisions that pit various policy goals against one another” (p. 5). Hence, SWB measures may help to evaluate the effects of previous policy decisions as well as assisting in formulating and predicting the effects of future policy alternatives (Diener et al., 2009).

Hence, it is essential that SWB measures are valid and reliable. Importantly, SWL may be measured using valid and reliable scales, such as the Satisfaction with Life Scale (SWLS; Diener et al., 1985). In an extensive review, Diener et al. (2013) present strong support for the ability of SWL measures to capture the quality of individual lives. They draw from a wide range of data showing that SWL scores: reveal significant differences between groups living in different circumstances; predict future behaviors, such as suicide, health and longevity; change in response to significant life events; and correlate with well-being measures that are not based on the respondent’s own self-report (for additional reviews, see Diener, Suh, Lucas, & Smith, 1999; Pavot & Diener, 1993). The strong support for the validity and reliability of SWL measures is indeed important if they are to be used in national progress indicators.

From a researcher’s perspective, it has been argued that the SWB approach with an SWL focus is *value neutral*, since it allows individuals to judge for themselves what they perceive as important for well-being (Diener et al., 1998). Diener et al. (2009) stress that: “[a]n advantage of subjective measures is that they reflect people’s desires and values, not just the judgments of the policy elites, and they are therefore inherently democratic in nature.” (p.47). This contrasts with a large number of other approaches related to well-being that comprise a fixed number of dimensions that are collectively said to represent individual well-being; for example, see the conceptualization of Ryff’s (1989) Psychological Well-Being (six dimensions, including autonomy, positive relations, self-acceptance, personal growth, purpose in life and environmental mastery) or Keyes’ (1998) Social Well-Being (five dimensions, including social integration, social coherence, social actualization, social contribution and social acceptance). As measures capturing Psychological Well-Being or Social Well-Being clearly demarcate what well-being is, they are not seen as SWB measures.

However, it may be argued that one-sidedly operationalizing the cognitive component of SWB as SWL is in and of itself an act of imposing values (e.g., see Christopher, 1999; Kjell, 2011; Kjell et al., 2016). That is, individuals may decide

what they think is important for their SWL; however, it is assumed that these individuals find SWL to be the most important aspect of well-being in the first place. Christopher (1999) argues that the SWL focus reflects the liberal individualism central to Western societies. In contrast, many Eastern cultures and traditions cherish harmony more than satisfaction (Joshani, 2014), where harmony is seen as a highly valued ideal (Li, 2006, 2008a). This questions the assumption that individuals value SWL as *the* overarching well-being construct, thereby suggesting that HIL may meaningfully complement the SWL focus within the SWB approach.

## HIL Meaningfully Complements SWL

To appreciate how HIL may complement SWL in a context of national progress, we first demonstrate how an HIL focus addresses problems associated with a mere SWL focus. Based on empirical findings, three interrelated reasons relevant for national progress are discussed. These include that an HIL focus: i) contributes to a comprehensive and representative view on well-being, rather than one that is narrow and unrepresentative; ii) encourages selflessness rather than just self-centeredness; and iii) accounts for the importance of interconnectedness for personal well-being.

### **A comprehensive and representative view of well-being**

Empirical evidence suggests that individuals do not primarily conceive well-being in terms of SWL, but rather as a mixture of constructs where HIL plays a key role. In a comprehensive study, Delle Fave et al. (2011) asked 666 participants from seven different countries to answer the question: *What is happiness for you?* Only 7.2% of the content was related to satisfaction, whereas 25.4% (the largest category) involved harmony and psychological balance. In another study, involving 2,799 participants from twelve countries from a broad range of different cultures (such as Italy, India, South Africa and the United States), Delle Fave et al. (2016) found that harmony was the most frequent description of happiness in all countries included in the study, except for Canada. They further pointed out that these descriptions of harmony are closely reflected in the Harmony in Life Scale (HILS) recently developed and validated by Kjell et al. (2016).

Kjell et al. (2016) argued that the unprecedented focus on SWL reflects a narrow view of cognitive SWB. SWL focuses on the evaluative judgment concerning to what degree life circumstances match one's personal expectations (Diener et al., 2009). In contrast, HIL implies favorable relationships in various aspects of one's life (Li, 2006), encompassing psychological balance and flexibility in harmonizing the various aspects of life (Kjell et al., 2016). Fittingly, Kjell et al. (2016)



demonstrated that HIL and SWL, despite the fact that they are strongly correlated, form a two-factor model of cognitive SWB. When further comparing the two well-being constructs, HIL explained more unique variance in relation to psychological well-being. So, one might argue that only focusing on SWL in national progress indicators represents a limited perspective. This narrow view of human well-being warrants serious concern, as it is potentially associated with undesirable consequences.

HIL exhibits stronger negative correlations to mental health problems compared to SWL. Kjell et al. (2016) found that the HILS explains more unique variance compared to the SWLS with regard to depression, anxiety and stress as measured by the short version of the Depression, Anxiety and Stress Scales (Sinclair et al., 2012). Likewise, Kjell et al. (in revision) used the same scale and found that the HILS yielded stronger negative correlations in relation to all three constructs. In a separate study, it was found that the HILS yielded stronger negative correlations to depression and worry as measured by the Patient Health Questionnaire-9 (Kroenke & Spitzer, 2002) and the Generalized Anxiety Disorder Scale-7 (Spitzer, Kroenke, Williams, & Löwe, 2006), respectively. Further, Kjell et al. (in revision) developed and validated a new method for statistically measuring, differentiating and describing psychological constructs by means of questions enabling open-ended word responses that are then analyzed using statistical semantics. These new semantic measures also revealed that HIL yields stronger negative correlations in relation to both depression and worry compared to SWL. Consistently, and applying different measures and methodologies, HIL is associated with a lower degree of self-reported mental health problems compared to SWL. Overall, this suggests that HIL more strongly reflects a mind free of mental health problems than SWL.

### **Nurturing selflessness rather than competing self-interests**

SWL encourages individuals to evaluate how their circumstances and their expectations match. This creates a risk to foster self-centeredness and, from an interpersonal perspective, create competing self-interests within and between groups of people. Kjell (2011) argues that SWL evaluations encourage individuals to put their own expectations first. This one-sidedly measuring of satisfaction risks leading to competing self-interests. He highlights that “one person’s satisfaction can result in another person’s dissatisfaction” (p. 260). When comparing SWL and HIL by means of statistical semantics, Kjell et al. (2016) found that individuals associated their pursuit of SWL with words related to self-centeredness, self-interest, in addition to their independence, mastery and personal achievements. In contrast, individuals associated their pursuit of HIL with words related to selflessness, psychological balance and flexibility, as well as their sense of relatedness, interconnectedness and being at peace. Hence, over-emphasizing the

individuals' own expectations, rather than encouraging individuals to contextualize their judgements, nurtures self-centeredness and increases the competition of self-interests.

Increased self-centeredness and competing self-interests might be particularly alarming bearing in mind the plethora of evidence indicating the significance of interconnectedness for individual well-being. For example, in a comprehensive review, belongingness and interpersonal attachment are portrayed as strong and fundamental human motivations (Baumeister & Leary, 1995). Likewise, relatedness to other people is considered a basic psychological need for human well-being (Ryan & Deci, 2000). Although empirical evidence shows that SWL increases with higher degrees of perceived positive relationships with others, HIL exhibits a significantly stronger correlation with positive relationships with others (Kjell et al., 2016).

### **Addressing the importance of interconnectedness**

Other forms of interconnectedness are also found to be related to well-being. Zelenski and Nisbet (2014) assessed connectedness with an adaptation of Aron, Aron and Smollan's (1992) inclusion of the other in the self-measure, where several pairs of circles labelled *self* and *other* represent different levels of connectedness by varying to what degree they overlap. When Zelenski and Nisbet (2014) relabelled the circles *me* and *nature*, they found that the degree of subjective sense of connectedness to nature significantly relates to several forms of well-being, even when controlling for the individual's general sense of connectedness (e.g., by changing the circle labelled nature to family, friends, one's country or culture). Further, in a meta-review, Capaldi, Dopko, and Zelenski (2014) found a consistent correlation between connectedness with nature and various well-being measures. Whereas SWL demonstrated the weakest relationship among the well-being measures, HIL appears better at capturing this relationship. For example, consider the fact that the words *nature* and *unity* are significantly related to individuals' pursuit of HIL as compared with SWL (Kjell et al., 2016). Further, Kjell et al. (2016) measured a general sense of interconnectedness to the world as measured with Circles of Life, comprising pairs of circles labelled *self* and *world*. This was compared with an adaptation of Cantril's (1965) Ladder of Life, where individuals are asked to select a step on a ladder where the lowest step is labelled *the worst possible life* and the highest step is labelled *the best possible life*. It was argued that the Circles of Life represent interconnectedness, interdependence and harmony, whereas the Ladder of Life represents independence, self-enhancement, self-centeredness and satisfaction. According to the hypothesis, the HILS explained more unique variance in the Circles of Life compared to the Ladder of Life, whereas

the SWLS exhibited the opposite pattern. Hence, HIL complements SWL by capturing the value of interconnectedness in relation to well-being.

Although research employing SWB has resulted in numerous useful findings, we propose that the current one-sided focus on SWL represents a narrow take on human well-being. Evaluations based on SWL encourage judgments based on personal expectations and decontextualized self-interest, whereas HIL evaluations invite individuals to consider contexts and interconnectedness. Hence, on a national level, an over-emphasis on SWL might be coupled with risks associated with competing self-interests. It is thus important to consider how the differences between HIL and SWL relate to goals central to national progress, including economic, social, environmental and sustainability aspects.

## HIL and SWL as Indicators of National Progress

### Well-being and economics

Considering the fact that economic activity is afforded a central role in national progress indicators, it is important to consider how SWL and HIL relate to economics. To assume that SWL is *the* overarching well-being construct may actually reflect the individualistic assumptions found in economics. Traditional models in economics are criticized for assuming an overly simplistic, one-dimensional view of human nature by overemphasizing people's individualistic self-interest (e.g., see Kirchgässner, 2014; Sen, 1977; Tittenbrun, 2013). Individuals are thought of as only pursuing what is in their own personal interest rather than, for example, exhibiting other-regard and helping. Diener et al. (2009) write:

Ever since Aristotle, those who study well-being have recognized the importance of family, friends, and other forms of social contact. Despite this long intellectual history, economists and psychologists have tended over the past century to concentrate on *individual needs and aspirations*. Well-being has often been treated as an individual outcome that is based on *the pursuit and achievement of individual goals*. Both survey and experimental data on well-being, however, show the importance of the social context. Some of the most important factors that influence well-being revolve around *the social features of people's lives*. (p. 176, italics added)

Although Diener et al., (2009) acknowledge the fallacy of the self-interest assumption in both economics and psychology, one might argue that this is not sufficiently addressed with regard to the SWL focus. Based on our previous discussion, the individualistic assumption appears to permeate the current unprecedented focus on SWL. This is further illustrated when Diener et al. (2009)

state that their definition of well-being “is based on an individual’s own interests, needs, preferences, and desires, and is therefore similar to the concept of ‘utility’ in economics” (p. 9).

The words individuals use for describing their pursuit of SWL empirically demonstrate their focus on individualistic self-interests and lack of social features, such as other-regard. Kjell et al. (2016) found that compared with HIL, the words individuals use for describing their pursuit of SWL focus on individual needs (e.g., *job, education, work, food, house, fulfillment, car, money and wealth*), achievement of individual goals (e.g., *achievement, career, success and goals*) and self-involved experiences (e.g., *pleasure and gratification*). Further, Kjell et al. (in revision) found that individuals describe personal high levels of SWL as *happy, content, fulfilled, pleasure and gratified*. It is not argued that these aspects are unimportant for the well-being of individuals, but the authors rather point out the lack of words describing other-regard and, as Diener et al. (2009) phrased it, “the social features of people’s lives.”

From a perspective of national progress, an SWB focus predominantly measured as SWL has been argued to complement social and economic progress indicators (e.g., Diener et al., 2009). However, on its own, SWL largely appears to reflect aspects important in traditional economics models. It could be argued that this close link between SWL and economics contradicts the key reasons for developing a complement to economic indicators in the first place. Therefore, we next discuss how HIL reflects social and environmental contexts.

## **Well-being and society**

From a perspective of national progress, the pursuit of both HIL and SWL appears essential for a well-functioning society. In a book by OECD, Rychen and Salganik (2003) identify three categories of key competencies for individuals to master for achieving “a successful life *and* well-functioning society” (p. 104, italics added); these three categories of competencies relate to both HIL and SWL. The first category refers to the ability to interact in heterogeneous groups and consists of three subcategories (Rychen & Salganik, 2003). These three subcategories closely reflect words used by individuals for describing their pursuit of harmony as compared with their pursuit of satisfaction (Kjell et al., 2016). The subcategories include relating well to others (consider descriptive words such as *friendship, understanding and sympathy*), cooperating (e.g., *cooperation, together and unity*) as well as managing and resolving conflicts (e.g., *agreement, forgiveness and tolerant*). Further, how individuals describe a high level of HIL is related to similar words, such as *amicable, consensus and empathy* (Kjell et al., in revision). Hence, both the pursuit

and the experience of HIL are intimately linked to competencies essential for a well-functioning society and a successful life.

The second category is linked to the pursuit of both HIL and SWL by referring to competencies of “acting autonomously” (Rychen & Salganik, 2003, p. 90). It comprises competencies that empower individuals to make meaningful life plans by stressing the ability “to defend and assert one’s rights, interests, limits, and needs” (Rychen & Salganik, 2003, p. 96). These aspects appear to relate to a focus on SWL; for example, consider participant generated descriptive words such as *goal* and needs such as *food* and *house* (Kjell et al., 2016). However, it is worth pointing out that the category firmly relates to an HIL focus by stressing that acting autonomously should involve responsible considerations of a broader social context.

In addition, HIL has been related to empowering aspects of individuals’ lives. Even though some might critically argue that an HIL focus inevitably implies the weakening of a person’s aims, needs and sense of independence, Kjell et al. (2016) found that compared with SWL, HIL correlates significantly stronger with the majority of empowering aspects they studied. Compared with SWL, HIL exhibited a stronger correlation with an independent self-construal, as well as several of the dimensions of Ryff’s (1989) Psychological Well-Being Scales, including personal growth, purpose in life and environmental mastery. Furthermore, two studies carried out by Vainio and Daukantaitė (2015) revealed that gritty individuals may pursue their goals whilst simultaneously report high HIL. Their analyses further revealed that authenticity and a sense of coherence mediate the relationship between individual levels of grit and HIL (which was also the case for grit and SWL). They concluded that in order for grit to relate to HIL, it is essential that personal meaning is embedded in the pursuits of goals. Hence, empirical studies to date suggest that an HIL focus does not imply giving up on personal aims or plans. On the contrary, HIL has so far been related to several empowering aspects of life, which relate to acting autonomously.

The last category highlights the competency of “using tools interactively” (p. 97), which chiefly refers to the ability to efficiently understand and use socio-cultural tools, such as information, knowledge, language and computers (Rychen & Salganik, 2003). This category is mostly reflected in the words individuals use for describing their pursuit of SWL, including words such as *education*, *achievement* and *work* (Kjell et al., 2016), but perhaps also in *understanding*, which is significantly related to the pursuit of HIL. Thus, HIL and SWL complement each other in reflecting key aspects of well-being that relate to a successful life and a well-functioning society. An HIL focus adds relational and cooperative aspects of well-being within a broad societal context, which certainly appears essential for national progress.

## Well-being and the environment

An SWL focus risks commodifying the environment, whereas an HIL focus potentially encourages individuals to develop a caring connection with nature. Alarming, Winter (2000) states that systematic exploitation and pollution of natural resources are largely due to human behaviors and associated thoughts, values, attitudes and feelings. One of the main underlying reasons is attributed to an existing human-nature relationship commodifying nature. For example, in traditional economic models, nature is predominantly seen as a utility resource with a focus on its instrumental values (Gómez-Baggethun, De Groot, Lomas, & Montes, 2010; Kosoy & Corbera, 2010). Correspondingly, Kjell (2011) argues that traditional well-being approaches also commodify nature; for example, the SWL focus encourages individuals to alter and manipulate the environment according to their needs and expectations. In this mindset, nature exists to fulfill our physical needs and desires, whilst we convey little concern for other aspects than oneself (e.g., consider the descriptive words related to pursuing SWL in Kjell et al. (2016): *fulfilled, gratification, house, food, money and car*). However, nature can also be a source of well-being, inspiration and recreation in and of itself. In three studies, Mayer, Frantz, Bruehlman-Senecal, and Dolliver (2009) found that participants exposed to nature, as compared with an urban setting, reported greater connectedness to nature, increased positive emotions and higher capability to reflect on a life problem. Their analyses further revealed that being connected to nature was a significant mediator for the other two beneficial effects. Nature encompasses values beyond a mere needs-and-desires satisfaction for human well-being.

When measuring SWB for national progress, it appears to be particularly important to capture the versatile significance of nature, and as discussed above, an HIL focus appears particularly fitting for capturing individuals' connection to nature. To reiterate and elaborate, Kjell et al. (2016) found that the pursuit of HIL, as compared with SWL, is associated with the word *nature*, as well as concepts suggesting a mindset more connected to the environment, such as *balance, unity, accord* and *concord*. Kjell et al. (in revision) also found that these kinds of descriptive words are also indicative of reports of a high degree of HIL. This is particularly important from a pro-environmental perspective, considering that Schultz (2001) and Schultz, Shriver, Tabanico, and Khazian (2004) demonstrate that individuals who implicitly and explicitly display a higher level of being connected to nature report more concern for the biosphere (i.e., concerns for all living things). In addition, Mayer and Frantz (2004) found that being connected to nature is positively related to self-reported eco-friendly behaviors. Whereas it appears as though SWL fails to graciously embrace the intrinsic values of nature and the aspects of a meaningful human-nature relationship, an HIL focus appears particularly apt when it comes to

increasing the level of connection with nature and potentially protective behaviors towards nature.

## **The sustainability of well-being**

We have discussed how HIL and SWL are linked to economic, social and environmental dimensions of national progress. These three dimensions together relate to sustainability, which is an important aspect when developing alternative indicators to economic measures. Sustainability refers to “meet[ing] the needs of the present without compromising the ability of future generations to meet their own needs” (World Commission on Environment and Development, 1987, p .24). The essence of sustainability involves reconciliation and integration between environmental, social and economic concerns and considerations (to more fully appreciate the complexity of sustainability e.g., see Gibson, 2006; Lozano, 2008). It involves acknowledging the interdependence between ecological and human systems (Gibson, 2006). This, for example, involves the reconciliation of conserving the capacity of the environment to absorb the high level of stress caused by human activity and interests, such as individual health, safety and the pursuit of well-being, as well as economic interests, such as equitability and development (e.g., see Kajikawa, 2008).

Stiglitz et al. (2010) emphasize that well-being measures in national progress indicators should be “put in a context of sustainability” (p. 12). Diener et al. (2009) also appear to appreciate the importance of sustainability, stating that “[o]ne of the most pressing policy concerns in the world is the health of the environment, and the problems for the environment caused by economic development and population growth” (p. 148). However, Kjell (2011) points out that the focus on SWL reflects and reinforces “the predominant view of seeing the individual first and the group second, as well as the view of nature as a sole commodity rather than also including intrinsic values” (p. 263). However, it is worth noting that we have demonstrated how HIL, as compared with SWL, might beneficially relate to the three dimensions characterizing sustainability. Within the economic dimension, an SWL focus appears to correspond to the status quo of emphasizing self-interest, whereas an HIL focus to a greater extent links the pursuit of well-being to social and environmental considerations. Within the social dimension, an HIL focus is an essential key aspect of a well-functioning society emphasizing the importance of cooperating with and relating to others. Within the environmental dimensions, an HIL focus acknowledges the intrinsic values of nature rather than commodifying nature’s instrumental resources to recklessly meet desires and expectations.

Further, Kjell (2011) argues that HIL corresponds to the conceptual nature of sustainability, whereas there is a certain degree of tension with regard to SWL. The

focus on independence and self-interest in SWL does not account for the *interdependencies* essential for sustainability and embedded in the very nature of HIL. Li (2008b) states that “harmony is by its very nature relational. It is through mutual support and mutual dependence that things flourish” (p. 427). Accordingly, consider the descriptive words reflecting interdependence and interconnectedness that Kjell et al. (2016) found significantly related to individual pursuits of HIL, as compared with SWL: *balance, unity, understanding, accord, agreement, concord, together, consistency* and *nature*. Similar words are indicative of high levels of HIL (Kjell et al., in revision). Whereas SWL is said to be isolating and decontextualized, HIL is demonstrated to have a lot in common with the essential aspects of sustainability, potentially acting as a psychological link to sustainable living. HIL and sustainability share an emphasis on balance and interconnectedness.

## Concluding Remarks

Subjective measures are just that: *subjective*. Just because a person reports a high level of HIL does not necessarily mean that he or she enjoys an objectively more harmonious life compared to someone reporting a lower level of HIL. Likewise, a high level of subjectively reported HIL in a nation does not necessarily mean that they live more sustainably, just as a high level of reported SWL does not necessarily mean that more needs and desires actually are fulfilled. Therefore, it is obviously valuable to continue using objective indicators as well. Diener and Suh (1997) point out that as most objective national progress indicators indirectly measure people’s well-being, subjective national progress indicators offer additional information suitable for assessing findings from objective measures. Importantly, they state “[i]f objective and subjective indicators converge, the researcher can make more definitive conclusions about quality of life. Where objective and subjective measures diverge, a deeper analysis of the meaning of the indicators is required.” (p. 205). Similarly, when measures of HIL and SWL diverge, a deeper analysis is required, where it is particularly interesting to consider the differences in causes as well as short- and long-term consequences related to an HIL versus an SWL focus.

We have presented evidence showing that SWL appears to specifically capture the fulfillment of self-concerned needs and goals, whereas HIL captures psychological balance, interconnectedness and the contextual, relational qualities of social and environmental aspects. We do *not* argue that SWL is of little importance for individuals, but that HIL is (at least) as important. A large number of studies demonstrate the importance of SWL (e.g., see Diener et al., 2013; Diener et al., 1999) and the SWL focus in research has resulted in vital progress in terms of our understanding of human well-being. Although HIL has a long-standing history



within philosophy (e.g., see Li, 2008a, 2008b), empirical research has indeed been conspicuous by its absence. Advancing research on HIL will allow us to form a more comprehensive understanding of well-being. We propose that evaluating national progress in part conceptualized as HIL is important, considering its potential and wide-ranging importance not only on an individual level, but potentially also for the environment and society at large.

## References

- Abdallah, S., Thompson, S., Michaelson, J., Marks, N., & Steuer, N. (2009). The Happy Planet Index 2.0: Why good lives don't have to cost the Earth.
- Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, 63(4), 596-612.
- Baumeister, R. F., & Leary, M. R. (1995). The need to belong - desire for interpersonal attachments as a fundamental human-motivation. *Psychological Bulletin*, 117(3), 497-529.
- Cantril, H. (1965). *The Pattern of Human Concerns*. New Brunswick: Rutgers University Press.
- Capaldi, C. A., Dopko, R. L., & Zelenski, J. M. (2014). The relationship between nature connectedness and happiness: a meta-analysis. *Frontiers in psychology*, 5. doi:10.3389/fpsyg.2014.00976
- Christopher, J. C. (1999). Situating Psychological Well-Being: Exploring the Cultural Roots of Its Theory and Research. *Journal of Counseling & Development*, 77(2), 141-152.
- Costanza, R., Erickson, J., Fligger, K., Adams, A., Adams, C., Altschuler, B., . . . Williams, L. (2004). Estimates of the Genuine Progress Indicator (GPI) for Vermont, Chittenden County and Burlington, from 1950 to 2000. *Ecological Economics*, 51(1-2), 139-155. doi:<http://dx.doi.org/10.1016/j.ecolecon.2004.04.009>
- Costanza, R., Hart, M., Talberth, J., & Posner, S. (2009). Beyond GDP: The need for new measures of progress.
- Delle Fave, A., Brdar, I., Freire, T., Vella-Brodrick, D., & Wissing, M. P. (2011). The Eudaimonic and Hedonic Components of Happiness: Qualitative and Quantitative Findings. *Social Indicators Research*, 100(2), 185-207. doi:10.1007/s11205-010-9632-5
- Delle Fave, A., Brdar, I., Wissing, M. P., Araujo, U., Solano, A. C., Freire, T., . . . Nafstad, H. E. (2016). Lay definitions of happiness across nations: The primacy of inner harmony and relational connectedness. *Frontiers in psychology*, 7.
- Diener, E. (2000). Subjective well-being: The science of happiness and a proposal for a national index. *American Psychologist*, 55(1), 34-43. doi:10.1037/0003-066X.55.1.34
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, 49(1), 71-75.

- Diener, E., Inglehart, R., & Tay, L. (2013). Theory and Validity of Life Satisfaction Scales. *Social Indicators Research*, 112(3), 497-527.
- Diener, E., Lucas, E. R., Schimmack, U., & Helliwell, F. J. (2009). *Well-being for public policy*: Oxford University Press, USA.
- Diener, E., Oishi, S., & Lucas, R. E. (2015). National accounts of subjective well-being. *American Psychologist*, 70(3), 234.
- Diener, E., Sapyta, J. J., & Suh, E. (1998). Subjective Well-Being Is Essential to Well-Being. *Psychological Inquiry*, 9(1), 33-37. doi:10.1207/s15327965pli0901\_3
- Diener, E., & Seligman, M. E. P. (2004). Beyond Money. *Psychological Science in the Public Interest (Wiley-Blackwell)*, 5(1), 1-31. doi:10.1111/j.0963-7214.2004.00501001.x
- Diener, E., & Suh, E. (1997). Measuring quality of life: Economic, social, and subjective indicators. *Social Indicators Research*, 40(1-2), 189-216.
- Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin*, 125(2), 276-302.
- Dolan, P., & White, M. P. (2007). How can measures of subjective well-being be used to inform public policy? *Perspectives on Psychological Science*, 2(1), 71-85.
- Gibson, R. B. (2006). Beyond the pillars: sustainability assessment as a framework for effective integration of social, economic and ecological considerations in significant decision-making. *Journal of Environmental Assessment Policy and Management*, 8(03), 259-280.
- Gómez-Baggethun, E., De Groot, R., Lomas, P. L., & Montes, C. (2010). The history of ecosystem services in economic theory and practice: from early notions to markets and payment schemes. *Ecological Economics*, 69(6), 1209-1218.
- Goossens, Y., & Mäkipää, A. (2007). Alternative progress indicators to Gross Domestic Product (GDP) as a means towards sustainable development. *Beyond GDP*, 305.
- Joshanloo, M. (2014). Eastern conceptualizations of happiness: Fundamental differences with western views. *Journal of Happiness Studies*, 15(2), 475-493.
- Kajikawa, Y. (2008). Research core and framework of sustainability science. *Sustainability Science*, 3(2), 215-239.
- Keyes, C. L. M. (1998). Social well-being. *Social psychology quarterly*, 121-140.
- Kirchgässner, G. (2014). On self-interest and greed. *Journal of Business Economics*, 84(9), 1191-1209.
- Kjell, O. N. E. (2011). Sustainable Well-Being: A Potential Synergy Between Sustainability and Well-Being Research. *Review of General Psychology*, 15(3), 255-266. doi:10.1037/a0024603
- Kjell, O. N. E., Daukantaitė, D., Hefferon, K., & Sikström, S. (2016). The Harmony in Life Scale Complements the Satisfaction with Life Scale: Expanding the Conceptualization of the Cognitive Component of Subjective Well-Being. *Social Indicators Research*, 126(2), 893-919. doi:10.1007/s11205-015-0903-z
- Kjell, O. N. E., Kjell, K., Garcia, D., & Sikström, S. (in revision at *Psychological Methods*). Semantic Measures: Using Natural Language Processing to Measure, Differentiate and Describe Psychological Constructs.

- Kosoy, N., & Corbera, E. (2010). Payments for ecosystem services as commodity fetishism. *Ecological Economics*, 69(6), 1228-1236.
- Kroenke, K., & Spitzer, R. L. (2002). The PHQ-9: a new depression diagnostic and severity measure. *Psychiatr Ann*, 32(9), 1-7.
- Li, C. (2006). The confucian ideal of harmony. *Philosophy East & West*, 56(4), 583-+. doi:10.1353/pew.2006.0055
- Li, C. (2008a). The Ideal of Harmony in Ancient Chinese and Greek Philosophy. *Dao*, 7(1), 81-98.
- Li, C. (2008b). The Philosophy of Harmony in Classical Confucianism *Philosophy Compass*, 3(3), 13.
- Lozano, R. (2008). Envisioning sustainability three-dimensionally. *Journal of cleaner production*, 16(17), 1838-1846.
- Mayer, F. S., & Frantz, C. M. (2004). The connectedness to nature scale: A measure of individuals' feeling in community with nature. *Journal of Environmental Psychology*, 24(4), 503-515.
- Mayer, F. S., Frantz, C. M., Bruehlman-Senecal, E., & Dolliver, K. (2009). Why is nature beneficial?: The role of connectedness to nature. *Environment and Behavior*, 41(5), 36.
- Pavot, W., & Diener, E. (1993). Review of the Satisfaction With Life Scale. *Psychological Assessment*, 5(2), 164-172. doi:10.1037/1040-3590.5.2.164
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68-78.
- Rychen, D. S., & Salganik, L. H. (2003). *Key competencies for a successful life and well-functioning society*: Hogrefe Publishing.
- Ryff, C. D. (1989). Happiness is everything, or is it - explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, 57(6), 1069-1081.
- Schultz, P. W. (2001). The structure of environmental concern: Concern for self, other people, and the biosphere. *Journal of Environmental Psychology*, 21(4), 327-339.
- Schultz, P. W., Shriver, C., Tabanico, J. J., & Khazian, A. M. (2004). Implicit connections with nature. *Journal of Environmental Psychology*, 24(1), 31-42.
- Sen, A. K. (1977). Rational fools: A critique of the behavioral foundations of economic theory. *Philosophy & Public Affairs*, 317-344.
- Sinclair, S. J., Siefert, C. J., Slavin-Mulford, J. M., Stein, M. B., Renna, M., & Blais, M. A. (2012). Psychometric evaluation and normative data for the Depression, Anxiety, and Stress Scales-21 (DASS-21) in a nonclinical sample of U.S. adults. *Evaluation & the Health Professions*, 35(3), 259-279. doi:10.1177/0163278711424282
- Spitzer, R. L., Kroenke, K., Williams, J. W. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The gad-7. *Archives of Internal Medicine*, 166(10), 1092-1097. doi:10.1001/archinte.166.10.1092
- Stiglitz, J. E., Sen, A., & Fitoussi, J.-P. (2010). Report by the commission on the measurement of economic performance and social progress. *Paris: Commission on the Measurement of Economic Performance and Social Progress*.

- Talberth, J., Cobb, C., & Slattery, N. (2007). The genuine progress indicator 2006. *Oakland, CA: Redefining Progress*, 26.
- Tittenbrun, J. (2013). The death of the economic man. *International Letters of Social and Humanistic Sciences*(11), 10-34.
- UNDP. (2015). *Human Development Report 2015: Work for human development*. Retrieved from Online: [http://hdr.undp.org/sites/default/files/2015\\_human\\_development\\_report.pdf](http://hdr.undp.org/sites/default/files/2015_human_development_report.pdf)
- Vainio, M. M., & Daukantaitė, D. (2015). Grit and Different Aspects of Well-Being: Direct and Indirect Relationships via Sense of Coherence and Authenticity. *Journal of Happiness Studies*, 1-29.
- Winter, D. D. N. (2000). Some big ideas for some big problems. *American Psychologist*, 55(5), 516.
- World Commission on Environment and Development (WCED) (1987). *Our common future*. United Nations.
- Zelenski, J. M., & Nisbet, E. K. (2014). Happiness and feeling connected the distinct role of nature relatedness. *Environment and Behavior*, 46(1), 3-23.



# Perspective II

## Statistical Semantics: Measuring and Describing Psychological Phenomena through Natural Language\*

### Abstract

The words and language used by individuals for expressing themselves contain a wealth of quantifiable information. Hence, we propose that statistical semantics, which quantify the meaning of words to enable statistical analyzes, have the potential of offering broad contributions in empirical research. Techniques within statistical semantics, and related methods from artificial intelligence, natural language processing and machine learning, are useful tools that may improve the quality of psychological research. We here describe common types of automated text analyses focusing on Latent Semantic Analysis (LSA), which is compared to the more traditional word frequency strategy referred to as Linguistic Inquiry and Word Count (LIWC). We further exemplify how statistical semantics has been applied within psychological research. We focus on three broad areas: i) self-reports data, where semantic measures based on statistical semantics may measure, describe and differentiate between psychological constructs; ii) naturally occurring (big) data, where statistical semantics may be applied for analyzing psychological aspects of social media texts, emails, blogs, etc.; and iii) enhancing experimental control and manipulations. We propose that statistical semantics offers great possibilities considering it is a flexible research tool with objective, systematic and quantitative qualities, where accrued research supports the validity and reliability of its application.

*Keywords:* Psychological assessments; Statistical semantics; Artificial intelligence; Natural language processing; Machine learning; Latent semantic analysis; Semantic measures; Vector space models; Linguistic analysis; Linguistic Inquiry and Word Count (LIWC).

---

\* This perspective will be submitted for potential publication with Katarina Kjell and Sverker Sikström as co-authors.

Language is the most common and reliable way for people to translate their internal thoughts and emotions into a form that others can understand. Words and language, then, are the very stuff of psychology and communication.

(Tausczik & Pennebaker, 2010, p. 25)

How individuals express themselves with words constitutes rich information that is fundamental for a wide range of empirical research. Typically, text data has been viewed as qualitative rather than quantitative data; however, we argue that this classification depends on which method is used for analyzing this data. We propose that text data may and should increasingly be collected and analyzed by means of quantitative methods. In science, quantification is fundamental for progress. Hence, precise and sophisticated methods that quantify and statistically analyze the meaning of words have the potential of offering an extensive contribution as a research tool within psychological science. In 2003, Pennebaker and colleagues reviewed a variety of successful uses of automated text analyses in psychology, which at that time mostly concerned conventional word frequency techniques involving counting the number of words belonging to predefined word categories. However, developments in artificial intelligence, natural language processing and machine learning have now enabled a variety of techniques we here argue have the potential of further improving psychological science. For example, Kjell, Kjell, Garcia, and Sikström (in revision) have developed semantic measures, where the open-ended word responses of individuals are analyzed using statistical semantics. In effect, this has important methodological implications, as it enables individuals to freely describe their mental states rather than being confined to closed-ended numerical rating scales. As discussed further below, their results demonstrated that semantic measures exhibited a higher or competitive level of reliability and validity compared to corresponding rating scales, but also that semantic measures complement and extend rating scales. This is because semantic measures appear to better differentiate between psychological constructs than rating scales; where the word responses describe the constructs. Thus, the proposed method measures, differentiates and describes the to-be-measured constructs.

This perspective is organized as follows: First, we review different types of text analysis approaches, including word frequency analyses (focusing on Linguistic Inquiry and Word Count [LIWC]; Pennebaker, Francis & Booth, 2001) and statistical semantics (focusing on Latent Semantic Analysis; LSA; Landauer & Dumais, 1997). These approaches are evaluated in terms of their objective, systematic and quantitative qualities. Second, we describe in more detail how statistical semantics based on LSA may be carried out for a variety of statistical analyses. Third, examples of how statistical semantics has been applied in psychological research are discussed. This concerns three broad areas: i) the self-reported accounts of individuals; ii) naturally occurring (big) data, and iii) within

experiments. We show that statistical semantics and associated methods unlock new possibilities in terms of analyzing data with the capacity to complement, and even extend, quantitative approaches traditionally used in psychology today. Finally, we elaborate upon the potential of statistical semantics, where we focus on: its qualities of being objective, systematic and quantitative; support for its validity and reliability; its flexibility in terms of type of data to collect and ways of analyzing this data; as well as its ability to replicate and extend previous research. Before we conclude, we also discuss potential limitations and challenges associated with statistical semantics.

## Types of Automated Text Analyses

Automated text analyses comprise a wide range of methods used for different purposes and associated with various strengths and weaknesses. Here, we discuss two broad approaches to automated text analyses that have been applied within various areas of psychology: word count strategies and statistical semantics (or word pattern analyses).

### **Word frequency strategies**

Word frequency strategies count the occurrence of words that have been categorized into various categories by the experimenter or independent judges. In psychology, one of the most commonly used programs is Pennebaker, Francis and Booth's (2001) LIWC (see also Stone and Hunt's (1963) General Inquirer and Hart's (2001) Diction). The LIWC program presents the results from a text as a percentage of words belonging to several predefined categories. LIWC2015 (Pennebaker, Boyd, Jordan, & Blackburn, 2015) facilitates comparisons of various texts in regards to more than 90 variables, such as pronouns (e.g., I, we), standard linguistic dimensions (e.g., common verbs, common nouns), psychological aspects (e.g., cognition, affect) and personal concern categories (e.g., work, leisure, home).

Counting words is a straightforward approach. However, it is ultimately a top-down approach relying on hand-coded categories based on the evaluations of judges or on word categories as defined by the experimenter. For example, the categorization of words in LIWC15 dictionaries regarding psychological constructs and personal concerns have been categorized by two to three judges. These judges have for instance evaluated whether or not a word pertains to positive emotions. Hence, in this respect the word categories are based on subjective interpretations (Pennebaker, Mehl, & Niederhoffer, 2003). This binary categorization also fails with regard to satisfactorily representing nuances and the complex relationships between words.



For instance, the category for anger in LIWC15 assigns kill and annoyed the same weight, which fails to represent differences in for instance valence, arousal and dominance. Further, the categorization procedure is costly and time-consuming, which means that for practical reasons, not all words are categorized. This naturally limits the analyses. Pennebaker et al. (2003) states:

Content-based dictionaries that are aimed at revealing what people are saying have not yielded particularly impressive results owing in large part to the almost infinite number of topics people may be dealing with. With the rapidly developing field of artificial intelligence, the most promising content or theme-based approaches to text analysis involve word pattern analyses such as LSA [i.e., Latent Semantic Analysis]. (p. 571)

### **Statistical semantics: Word pattern analyses**

Instead of using word categories constructed by experimenters or judges, data-driven bottom-up approaches may instead be used for representing meaning, or the semantics, of words. Commonly used unsupervised approaches include Latent Semantic Analysis (LSA; Landauer & Dumais, 1997), latent Dirichlet allocation (LDA; Blei, Ng, & Jordan, 2003) and word embeddings (e.g., see Mikolov, Sutskever, Chen, Corrado, & Dean, 2013). These are unsupervised, as the semantics are not derived from given categories/labels or judges. To derive the semantics of words, these approaches rely on the statistical patterns of word use and may broadly be referred to as statistical semantics. Turney and Pantel (2010) point out that the term statistical semantics was used by Furnas, Landauer, Gomez, and Dumais (1983) without offering a definition and was later described on Furnas' faculty webpage as the "studies of how the statistical patterns of human word usage can be used to figure out what people mean" (as cited in Turney & Pantel, 2010, p. 146; see also Weaver's (1955) use of the term). Statistical semantics is for instance studied within computational linguistics, natural language processing, artificial intelligence and cognitive science. These approaches are also broadly referred to as Vector Space Models (e.g., Turney & Pantel, 2010), Distributional Semantics Models (e.g., Lapesa & Evert, 2014), Topic Models (e.g., Atkins et al., 2012) or Probabilistic Topic Models (e.g., Blei, 2012). Next, we focus on describing LSA as this approach has been used in a variety of contexts in psychology. Even though we do not describe the related LDA and word embeddings approaches in detail, we present some research based on these approaches (for a review of LDA, see Blei, 2012).

"You shall know a word by the company it keeps" (Firth, 1957, p. 11) is a core rationale of word pattern analytical approaches such as LSA. Statistically speaking, the ways in which words are used within a language are not randomly distributed

(Iliev, Dehghani, & Sagi, 2014). Instead, the distribution of contextual words is predictable and defines its meaning. Hence, LSA is using this distribution to represent “the meaning of a word through the contexts in which it has been observed in a corpus” (Erk, 2012, p. 635). In practice, this is carried out by creating a matrix of word co-occurrence counts, which then undergoes an analysis similar to a factor analysis. For example, the rows of the frequency matrix contain the words in a language, the columns contain word contexts and the cells contain the co-occurrence counts/frequency. The dimension reduction method used for LSA is referred to as singular value decomposition (Golub & Kahan, 1965; Landauer & Dumais, 1997) and is akin to principal component analysis (Iliev et al., 2014). The factors are typically referred to as dimensions and around 300 to 800 dimensions are usually extracted. Ultimately, each word is represented by a vector, or a semantic representation, containing a number for each dimension. These values may be seen as the coordinates of a point in a high-dimensional semantic space. The closer two points are situated in this semantic space, the more similar they typically are in meaning. In other words, LSA captures the relationships between words, where proximity in the semantic space indicates semantic similarity.

Importantly, Landauer and Dumais (1997) stress that LSA-based word similarities are captured by means of indirect, higher order associations. Whereas first order co-occurrences capture the relationship between words that explicitly occur together in a context, higher order associations capture the relationship between words that do not necessarily appear in the same context, but share first order co-occurrences. For example, doctor and physician rarely co-occur in the same sentence (i.e., first order co-occurrence), whereas they both share contextual words such as hospital, nurse and disease (i.e., higher order co-occurrences). This is important, as Landauer, Foltz, and Laham (1998) state:

[T]he similarity estimates derived by LSA are not simple contiguity frequencies, co-occurrence counts, or correlations in usage, but depend on a powerful mathematical analysis that is capable of correctly inferring much deeper relations (thus the phrase latent semantic), and as a consequence, they are often much better predictors of human meaning-based judgments and performance than are the surface-level contingencies (p. 260-261).

### *The validity of LSA*

Psychologists often use factor analyses in order to, for example, investigate personality traits that may then be used for predicting behaviors. The factor solution is typically selected based on statistical properties and theoretical relevance. In LSA, on the other hand, the number of dimensions is typically selected based on the performance in some external validity test, such as a synonym test. The synonym test involves presenting a target word and a set of alternative words, where one of these alternative words are a synonym of the target word. To complete the test, one

selects the word with the highest LSA-generated semantic similarity to the target word. The optimal number of dimensions to extract is the dimension solution that correctly selects the most synonyms.

At the optimal number of extracted dimensions, Landauer and Dumais (1997) found that their LSA-based semantic space achieved 64.4 percent correct answers on the synonym part of a Test of English as a Foreign Language (TOEFL). This is remarkably close to the average rate of correct answers (64.5%) by the foreign students taking the test, where an average score is sufficient for being admitted to several universities. Rapp (2003) improved this procedure by among other things using raw data with shorter word contexts of two words before and after the target word rather than entire documents. This method achieved an impressive 92.5 percent correct answers on the same TOEFL.

LSA also performs results comparable with human raters on various semantic tasks, as well as corresponds to neural brain activity. Foltz, Laham, and Landauer (1999) found that essays graded by humans versus LSA-based methods correlate almost as strongly as between two human graders ( $r = .701$  and  $.707$ , respectively). They argue that LSA provides an objective, reliable and fast way of grading essays. Further, based on high performances on multiple validity tests/tasks, LSA is proposed to represent a computational model of how knowledge is represented in humans (Landauer & Dumais, 1997). Indeed, there is some data supporting a link between neural brain activity and LSA-based semantic representations. By means of brain activity recordings using functional magnetic resonance imaging (fMRI), Carlson, Simmons, Kriegeskorte, and Slevc (2014) found that LSA-generated semantic similarities of objects reflect both neural organization within the inferior temporal cortex (ITC; associated with late visual processes) as well as semantic similarity judgments made by individuals. They concluded that their “data suggest that measures like LSA do, in fact, reflect important aspects of how the human brain represents conceptual information.” (p. 129). Thus, overall word pattern analyses have performed well in multiple contexts.

## **Evaluation of word count strategies and statistical semantics**

In the context of content analysis, Berelson (1954) states that a text analysis seeks to be objective, systematic and quantitative. Objective denotes that the analysis should not rely on the subjective interpretations of the analyst, but should instead be unbiased in order to enable replications by independent researchers (Berelson, 1954). Both statistical semantics and word frequency strategies may be carried out so that the results are reliably reproduced by other researchers. As discussed, however, word frequency strategies such as LIWC may be seen as subjective in the construction of the actual categories, whereas statistical semantics is considerably

more data-driven. The only methodological decisions required when using such an approach include the type of word data to use when constructing the LSA-based semantic space or the type of validity test to use when deciding the number of dimensions to extract.

Systematic signifies that all parts of the texts should be analyzed methodically according to an explicit strategy (Berelson, 1954). Computer programs are by nature systematic, as they carry out tasks in accordance with their (explicit) code. Further, word frequency strategies might be seen as explicit considering the ease with which one may understand how words are counted and know the words composing the predefined categories. Whereas each individual semantic dimension created using LSA is typically difficult to interpret, we argue that this approach is more accurately modelling the complex and nuanced interrelationships among words. In addition, statistical semantics may represent an advantage as the semantic space comprises more words than are included in the categories of current word frequency strategies, which is a result of practical considerations.

Quantitative stresses procedures that enable counting and making statistical inferences, which is at the very core of the automated text analyses we have described. As discussed, word frequency strategies plainly count words as binary (i.e., either belonging to a category or not), whereas statistical semantics approaches are more nuanced. For example, LSA describes each word in a high dimensional space, where each word is typically described using 300 to 800 numbers.

Word frequency strategies and statistical semantics approaches currently share some unresolved limitations. First, sarcasm and irony are problematic to take into account. In this regard, it is assumed that most individuals, most of the time, actually mean what they say or write. Second, there are difficulties involved in managing words with several meanings (e.g., bank, which may refer to a river bank or a financial bank). However, this only applies to a rather small portion of all words and is thus not considered to pose a major threat. Finally, the discussed approaches treat the text as a “bag of words” by ignoring word order within a given text. The approaches cannot distinguish between “Daniele loves Victoria” and “Victoria loves Daniele”, but only represent the meaning as pertaining to love. However, there is research attempting to address this problem by incorporating word order into the model (e.g., see Jameel & Lam, 2013).

Even though both word frequency strategies and statistical semantics overall fare well in terms of the criteria of being objective, systematic and quantitative, one might argue that statistical semantics approaches live up to them more stringently. Furthermore, their shared limitations may not be considered major. Even though it is not possible to date to perform a complete extraction of meaning from text data, Turney and Pantel (2010) conclude that statistical semantics approaches are arguably “the most successful approach to semantics, so far” (p. 145). In addition,

as statistical semantics approaches extract meaning automatically from a given set of words, they require less effort compared to word count strategies such as LIWC, which require (several) judges to categorize every single word. In short, statistical semantics approaches offer great potential in terms of being used as objective, systematic and quantitative tools for analyzing semantic content in a variety of psychological settings.

## Using Semantic Representations in Analyses

Next, we demonstrate the breadth and flexibility of analyses based on statistical semantics. We describe some relevant analyses related to LSA and the use of semantic representations. For a clear description of the methodology involved in performing text analyses based on LDA using large amounts of social media texts, see Kern et al. (2016). All of the analyses described below may be carried out using [www.semanticexcel.com](http://www.semanticexcel.com) (a user-friendly, online software solution; see Sikström, Kjell & Kjell, Appendix B).

### **Semantic spaces and semantic representations**

The semantic representations for each word from the semantic space constitute the basis of using LSA, where semantic representations from single words may be added together to semantically represent longer texts, such as sentences or paragraphs. Producing high-quality semantic representations requires basing the semantic space on a massive amount of text data, which is rarely possible to collect in psychological research. However, it is possible to create a semantic space from other, unrelated data and then map the produced semantic representations to the words collected in a study. Whereas some researchers use big data to directly extract information and make inferences from this information (e.g., Schwartz et al., 2013), we use big data to assist us in analyzing small data sets. We refer to this approach as *small-by-big data analyses* (SBDA). For example, one may use informal diary-style writings by individuals from unrelated studies (e.g., Campbell & Pennebaker, 2003), news articles from online news outlets (e.g., Garcia & Sikström, 2013) or texts from published books (e.g., Kjell, Daukantaitė, Hefferon, & Sikström, 2016). Semantic Excel includes semantic spaces for several languages, including Chinese, Czech, Dutch, English, Finnish, French, German, Hebrew, Italian, Norwegian, Persian, Polish, Portuguese, Romanian, Russian, Spanish and Swedish.

Semantic spaces can also be specialized to capture different aspects of language. For example, frequently used “non-content” words such as *and*, *the* and *a* may be

excluded in order to enhance the semantic representations of content words. On the other hand, when examining stylistic aspects of language (i.e., *how* something is being said rather than *what* is being said), semantic spaces that only include certain aspects of a language, such as pronouns, prepositions or auxiliary verbs, may be used successfully (e.g., see Campbell & Pennebaker, 2003).

## Semantic similarity

The semantic similarity between two words/texts may be derived by computing the cosine of the angle between their semantic representations (Landauer & Dumais, 1997). The semantic representation may be seen as the coordinates for a point in a high-dimensional space, whereas the cosine of the angle between two points in this high-dimensional space indicates the relationship between the two points. The semantic similarity scores may be used for testing whether there is a significant difference between two sets of words/texts (for example see Kjell et al., in revision).

## Semantic-numeric correlations and semantic predictions using training

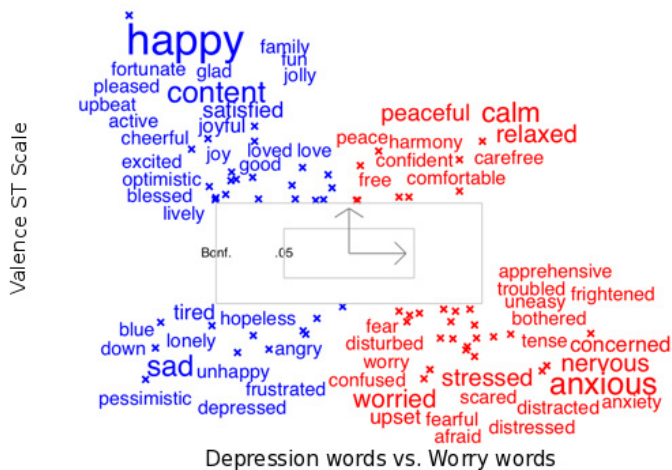
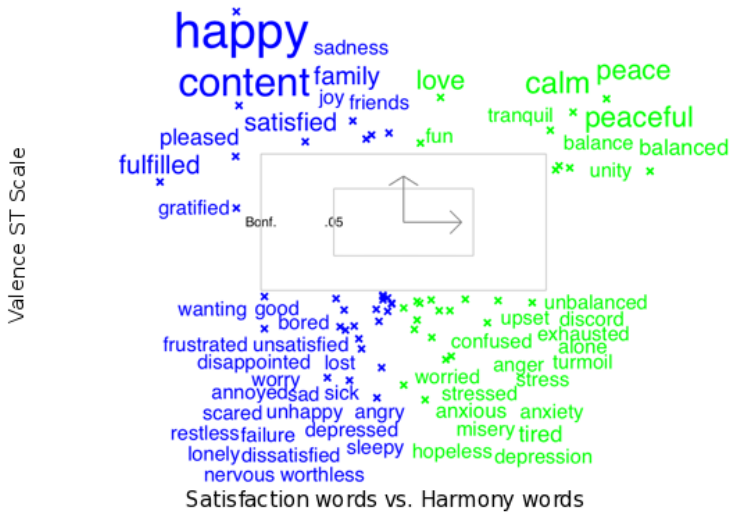
Using machine learning, the semantic representations may be used for analyzing the relationship between words/texts and numerical values, which we refer to as a *semantic-numeric correlation*. By using multiple regression ( $y = \beta_0 + \beta_1 * x_1 \dots \beta_m * x_m + \epsilon$ ), the semantic dimensions from the semantic representation ( $x_1$  through  $x_m$ ) may be used for predicting a numerical value ( $y$ ), such as objective measures, rating scales or categorical data (where  $\beta_0$  is the constant,  $\beta_1$  through  $\beta_m$  are the coefficients defining the relationship between the words and the numerical outcome and  $\epsilon$  is the error term). Using leave-n-out cross-validation, the relationship may be assessed by first training the model using a part of the data set (a training set) and then apply the coefficients ( $\beta_1$  through  $\beta_m$ ) to the dimensions of the semantic representation in the other part of the data set (the test set) in order to predict a value ( $\hat{y}$ ). The predicted values ( $\hat{y}$ ) are then correlated with the actual values ( $y$ ) to test the strength of the relationship between the words/texts and the numerical values (e.g., see Kjell et al., 2016; Kjell et al., in revision).

This approach lends itself to a high level of flexibility, as one may apply the coefficients ( $\beta_1$  through  $\beta_m$ ) from a previously trained model to another set of data, which we refer to as *semantic prediction*. In this manner, it is possible to study a large number of semantic features. For example, Semantic Excel includes already trained models for the Affective-Norms-for-English-Words (ANEW; Bradley & Lang, 1999), where individuals have rated words according to features, including valence, arousal and dominance. These models may now be applied to any words/texts in order to estimate the degree of these features. A definite strength of

using this method is that all words in the semantic space receive an estimate, as what is assigned coefficients are the dimensions of the semantic representations (i.e., it is not required that the word was evaluated in the original [ANEW] word list).

### **Describing the meaning of words through visualization**

Using keyword analyses in various ways highlight words that are significantly more represented in one of two groups of word data or according to specified dimensions within a data set (for an example, see Figure 1). For example, one may use chi-square tests in order to reveal words that differ significantly between two groups of texts. Or words within a data set may be plotted according to their correlation to a given scale being studied (such as an objective measure, a semantic prediction, a semantic similarity scale or a numerical rating scale). The difference in relevance between the words may be demonstrated further by representing their frequency by means of font size in a plot. As compared with mere numeric data, these kinds of keyword analyses certainly give a unique advantage when performing text analyses by descriptively visualizing the data (Kjell et al., in revision).



**Figure 1. Examples of what visualizations of keyword analyses may look like.**

The top plot compares individual word responses with the questions concerning their perceived satisfaction with life (left/blue) and harmony in life (right/green); the bottom plot compares individual responses with a question concerning depression (left/blue) and worry (right/red). The arrows show the origo for x- and y-axes; where the axes names are written in the outskirts, below and to the left, of the arrows. On the x-axes, words are plotted to the q-values from a Chi-square test comparing the responses to each question, using Bonferroni correction for multiple comparisons (Bonf. = Bonferroni line where  $q = 4.00$ , and  $.05$  indicates the uncorrected p-value line, where  $q = 1.96$ ). On the y-axes, the words are plotted according to their Pearson's point biserial correlation to the semantic predicted ANEW valence scale ( $r = .14$  at the Bonferroni line, and  $r = .07$  at the  $.05$  uncorrected p-value line). A larger font size indicates a higher frequency, using fixed lower and upper limits.  $N = 400$  participants from Study 9 by Kjell et al. (in revision) at time 2 (see their study for more details).



# Applying Statistical Semantics in Psychological Research

In psychology, statistical semantics may naturally be applied in research directly related to language; however, it may also be used as a statistical tool in order to analyze word data in a wider research context beyond language research. Landauer (1999) maintains that LSA captures “a theory of the psychology of language and mind” by arguing that “it offers a biologically and psychologically plausible mechanistic explanation of the acquisition, induction, and representation of verbal meaning” (p. 303). Naturally, statistical semantics has been used for studying cognitive aspects, such as semantic linguistic maturity in children (Hansson, Bååth, Löhndorf, Sahlén, & Sikström, 2015), episodic memory recall (Howard & Kahana, 2002), correctness of eyewitness statements (Sarwar, Sikström, Allwood, & Innes-Ker, 2015) and word association in patients diagnosed with Broca’s aphasia (Hansson et al., 2015). This is truly important, as it is an investigative tool directly related to understanding language and to further the understanding of meaning and knowledge representation. However, we argue that statistical semantics may also be applied as a broader research tool; for example, to analyze individuals’ open-ended reports regarding their state of mind, psychological aspects of naturally occurring texts and enhancing experimental control and stimuli.

## **Individuals’ accounts related to psychological phenomena**

### *Conceptualizing psychological constructs*

Understanding the meaning of constructs is a core objective in behavioral studies, as this defines what is actually being studied. Statistical semantics has successfully been employed in order to conceptualize psychological constructs. Kjell et al. (2016) have used various types of statistical semantics analyses for conceptualizing different well-being constructs. They asked participants mainly from India and United States to write ten words that describe how they pursue harmony in life, satisfaction with life, psychological well-being and happiness. By using semantic t-tests on the responses, they established that individuals view these concepts as significantly different and the effect sizes and the semantic similarity scores indicated that harmony in life is semantically closer to psychological well-being compared to satisfaction with life and happiness. Keyword analyses between harmony in life and satisfaction with life further revealed that these differed according to theoretically relevant aspects. The pursuit of harmony in life was significantly related to words linked to interconnectedness (e.g., *peace, balance, cooperation* and *agreement*), whereas the pursuit of satisfaction with life was related to words linked to independence (e.g., *job, money, achievement, education* and

*pleasure*). Employing semantic-numeric correlations, by training the semantic representations of harmony and satisfaction words to the respective rating scale, further revealed that how individuals describe their pursuits of harmony versus satisfaction is significantly related to their reported rating scale scores.

### *Measuring, differentiating and describing psychological constructs*

In contrast to the commonly used rating scales, where the construct is predefined by researchers constructing the scales, semantic measures enabling open-ended responses that are analyzed using statistical semantics may be used for describing the construct empirically. Kjell et al. (in revision) have demonstrated that the semantic measures approach may effectively measure as well as describe psychological constructs with competitive or higher levels of validity and reliability compared to closed-ended rating scales. They first investigated reports regarding external stimuli, where individuals described various facial expressions in pictures using either traditional rating scales or semantic questions with open-ended responses. Semantic predicted scales (i.e., training the semantic responses to the facial expressions), as well as semantic similarity scales between word responses and word norms that describe relevant facial expressions were compared with numerical rating scales. It was demonstrated that both semantic predicted scales and semantic similarity scales enabled a categorization of the facial expressions with a significantly higher level of accuracy compared to when using traditional numerical rating scales. It was also found that the semantic measures exhibited a significantly higher level of interrater reliability compared to rating scales in terms of both categorizing facial expressions accurately as well as related dimensions, including the valence, arousal, intensity, clarity and genuineness of the expressions.

In addition to the high level of validity and reliability in categorizing facial expressions, the semantic measures also include the advantage of describing facial expressions without priming respondents with the words otherwise necessary for defining the rating scales. Visualizations based on keyword analyses showed that happy facial expressions were most frequently described with the word *happy* and sad facial expressions were described with the word *sad*. Informatively, the keyword analyses also revealed that contemptuous facial expressions, which had been found difficult to categorize in previous research (see Langner et al., 2010), were frequently described as *annoyed* using semantic questions. Hence, the semantic measures approach may capture concepts associated with difficult semantic labels, as well as inform us about suitable descriptions of these concepts.

In a series of studies analyzing reports regarding subjective states, Kjell et al. (in revision) have demonstrated that semantic measures are capable of assessing subjective states of mind. Individuals were for instance asked to use descriptive words or a text for describing whether or not they experienced overall harmony in life. The validity of the semantic responses for harmony in life, satisfaction with

life, depression and worry were first demonstrated by showing that their semantic trained scales to the respective rating scale correlated strongly with the actual rating scale scores (for descriptive words:  $r = .58-.72$ ,  $p < .001$ ). Semantic ANEW valence predictions also correlated strongly with rating scale scores. Importantly, semantic similarity scales to word norms describing the constructs under investigation were also capable of measuring the psychological constructs completely independent from rating scales. In addition, semantic measures also demonstrated satisfactory test-retest reliability as well as lower social desirability as compared with rating scales.

By employing keyword analyses, semantic measures may also describe significant aspects of the states being studied. For example, individuals described harmony in life using words such as *peace*, *balance* and *agreement*; satisfaction with life using *happy*, *fulfilled* and *content*; depressed using *sad*, *lonely* and *blue*; and worried using *anxious*, *scared* and *nervous*. Hence, open-ended semantic questions analyzed by means of statistical semantics may both measure the degree of a psychological construct as well as describe it. In addition, Kjell et al. (in revision) found that their overall results suggest that rating scales tend to predominantly capture valence, whereas semantic similarity scales may better differentiate between constructs by capturing the targeted measures more clearly.

#### *Analyzing individual narratives*

Whereas rating scales require participants to explicitly rate the to-be-studied construct, statistical semantics allows studying a wealth of measures in narratives without making this explicit to the participants. This may include several advantages, such as avoiding revealing the research hypothesis to participants and increasing methodological flexibility. Further, using statistical semantics, one may study the relationship between individual personal narratives and psychological constructs and related behaviors. For example, it was found that when using trained semantic scales, the written descriptions concerning a positive or a negative life event could predict self-reported positive and negative affect in adolescents (Garcia & Sikström, 2013). Further, semantic-numeric correlations scales were used between the semantic representations of narratives concerning positive and negative life events by adolescents and rating scales measuring personality traits (Garcia, Anckarsäter, et al., 2015). It was found that self-directedness, cooperativeness and self-transcendence, but not the big five personality traits (i.e., openness, conscientiousness, extraversion, agreeableness and neuroticism), are involved when adolescents describe experiences relating to positive and negative life events. Karlsson, Sikström, and Willander (2013) have shown how statistical semantics provides us with a new way of studying the recall of personally experienced events (or autobiographical memories). They asked participants to narrate an autobiographical memory in response to visual, auditory, olfactory or multimodal

retrieval cues. They point out that previous studies on cued autobiographical memories have focused on when the memory occurred or on the rated subjective experiences of things such as valence. However, an important aspect in this regard is that with statistical semantics, the meaning is quantitatively analyzed. Semantic t-tests showed that the semantic content from the different texts differed significantly depending on the modality of retrieval cues.

Individual narratives have also been linked to health behaviors. Research has consistently demonstrated that individuals who write about an emotional upheaval, as compared to a non-emotional topic, and typically for 15–30 minutes on 3–5 consecutive days later exhibit better psychological and physical health (for a review, see Pennebaker & Seagal, 1999). Pennebaker, Mayne, and Francis (1997) revisited previous studies by applying word frequency analyses comparing the first with the last essays, which revealed that individuals who over the days of writing increased the number of words categorized as self-reflective (e.g., *understand*, *realize* and *consider*) and causal (e.g., *because*, *cause* and *reason*) exhibited a significant correlation to health improvements. They emphasized the importance of flexibility in writing/thinking patterns. An important aspect in this context is that when using LSA, Campbell and Pennebaker (2003) reanalyzed three studies analyzing the relationship between changes in writing and visits to a physician. Using a typical content word semantic space did not reveal any relationship; however, using a semantic space focused on style rather than meaning (i.e., only including particles in the space) revealed a strong relationship between individual writing and visits to a physician. Specifically, they found that individual flexibility in the use of pronouns (e.g., *me*, *we* and *us*) during the different days of writing (as measured using semantic similarity scores between days) exhibited a significant correlation with subsequent visits to a physician ( $r = .35\text{--}50$ ,  $p \leq .05$ ). They emphasized the potential of LSA, concluding that “[a]cross all three studies, the effect size was far greater than the effects we had found with any other analytic strategy” (p. 62).

## **Psychological aspects of naturally occurring (big) data**

### *Linking survey responses and naturally occurring data*

Combining survey responses and naturally occurring data may be a valuable technique for generating new insights with high ecological validity. Statistical semantics may enhance analyses of naturally occurring data, such as analyses of individual social media texts, emails, blogs, text messages, letters, diaries, transcribed events from real-world events (i.e., out of labs), transcribed or recorded therapy sessions and so on. In addition, these sources of information may be combined with self-reported measures. For example, survey responses were combined with naturally occurring texts by asking individuals to answer personality

inventories as well as to provide their fifteen most recent status updates on Facebook (Garcia & Sikström, 2014). This revealed that semantic-numeric correlations between status updates and personality rating scales had a significant correlation with psychopathy, neuroticism and narcissism, but also that psychopathy and narcissism were negatively correlated with semantic ANEW valence predictions.

### *The importance of big data*

As statistical semantics analyses are automated, they enable us to study really large data sets, or big data, of naturally occurring text, where manual coding is extremely time-consuming and frequently unreliable (for a review of big data, see Chen, Mao, & Liu, 2014). Further, it is typically unrealistic to collect participant-generated questionnaire data in such a quantity so as to categorize it as big data. However, big data is considered increasingly important for future research. Markowitz, Błaszczewicz, Montag, Switala, and Schlaepfer (2014) suggest that it will be more common to study and perform deeper analyses on already existing big data rather than designing studies. They hypothesize that big data and related technologies will be of great importance for psychometric research and applied settings. Lazer et al. (2009) point out that “a computational social science is emerging that leverages the capacity to collect and analyze data with an unprecedented breadth and depth and scale.” (p. 722). Kosinski, Wang, Lakkaraju, and Leskovec (2016) point out that big data samples are capable of revealing patterns that are difficult to find in small samples and provide high statistical power. Hence, automated text analysis is a key component in the growing interest in big data.

Tailored to the needs of big data analyses, Schwartz et al. (2013) have developed the Differential Language Analysis (DLA) to distinguish language features based on words, phrases and latent Dirichlet allocation-derived topics. DLA focuses on finding language features that strongly correlate with a variable under investigation (e.g., personality traits, age). Analyzing millions of Facebook messages from 75,000 individuals, they were able to distinguish between gender, age and the big five personality traits as measured by rating scales. They furthermore found that their method outperformed models based on the LIWC word frequency approach in predicting these attributes. Their models predicted self-reported personality traits with correlations ranging from  $r = .31-.41$ , whereas the LIWC-based analyses ranged from  $r = .21-.29$ . Park et al. (2015) further examined the validity and reliability of these language-based assessment models. For example, they found that the measures agreed with both self-reported as well as the informant’s account of personality and that the measures were stable over a six-month period. Schwartz et al. (2016) have also used texts from social media for predicting satisfaction with life. Lastly, Eichstaedt et al. (2015) have used text from tweets in order to predict heart disease mortality on a county level, where they among other language features used LDA-derived topics. Importantly, Schwartz et al. (2013) conclude that:

Over the past one-hundred years, surveys and questionnaires have illuminated our understanding of people. We suggest that new multipurpose instruments such as DLA emerging from the field of computational social science shed new light on psychosocial phenomena. (p.14)

As an example of systematically analyzing large and naturally occurring data using LSA, Gustafsson Sendén, Lindholm, and Sikström (2014) studied self- and group-serving biases by examining the use of pronouns (*I, we, he, she* and *they*) in news texts. They created a semantic space out of 800,000 news messages, after which they extracted pronoun word contexts by gathering 15 words before and after each pronoun from 50 percent of the articles. Using semantic predictions of ANEW valence on these word contexts, they found that personal pronouns that included oneself (i.e., *I* and *we*) occurred in contexts that were significantly more positive in valence compared to pronouns that exclude oneself (i.e., *he, she* and *they*). Gustafsson Sendén, Sikström, and Lindholm (2015) furthermore found that the contexts of *he* were more positive in valence as compared with contexts of *she*. In both studies, the authors point out that the valence-related finding was of a small effect size. However, considering the wide, daily distribution of these news, the authors propose that they might have a considerable impact on reinforcing and contributing to the revealed biases (see also Garcia, Kjell, & Sikström, 2015 in relation to big samples of news article and happiness).

## **Enhancing experimental control and effects**

Statistical semantics may be used for constructing and examining stimuli for experiments. For example, for a recognition task Dougal and Rotello (2007) constructed positive, negative and neutral word stimuli that were matched in LSA-based semantic similarity. They drew attention to the importance of semantic aspects by concluding that the accuracy of recognition is not better with regard to emotion words when semantic similarity is controlled for across the word stimuli. Similarly, Gagné, Spalding, and Ji (2005) used LSA in order to examine word stimuli used in previous experiments on relational priming and found that semantic similarity systematically differed across conditions. When they controlled for semantic similarity, the effect of relational information disappeared. Hence, statistical semantics may be used for improving the control of stimuli within experiments.

Statistical semantics may also be used for studying the effects of language-based manipulations. A fairly common manipulation in psychology involves asking participants to write about different topics, such as autobiographical recall; for example, to induce a specific emotion (e.g., see Lench, Flores, & Bench, 2011). Normally, the written texts are not analyzed. Instead, participants from different

conditions are compared in relation to an outcome variable. It is however worth noting that as a control measure, semantic t-tests may help assess whether different conditions in fact resulted in different semantic content, and the effect size indicates how well the manipulation worked. One may also analyze whether the text differs in semantic similarity to specific word norms or in semantic predicted scales, such as valence. Keyword analyses may furthermore reveal important aspects both between and within conditions. Lastly, one may also use semantic predicted condition scales (i.e., training semantic content to the type of condition) in order to select the participants who were most responsive to the manipulation (Garcia & Sikström, 2013). In other words, comparing participants in different conditions focuses on comparing the different manipulation instructions, whereas selecting participants based on semantic predicted condition scales (i.e. only including those where the prediction may clearly distinguish between conditions using a set cutoff point) focuses on the actual behaviors elicited in response to each condition. Garcia and Sikström (2013) point out that the former may be seen as an indirect measure and the latter as a direct measure of behavior, as it more directly examines the behavior being studied. Using this method, they magnify the relationship between participants' self-reported affect and texts regarding a positive or negative event. Hence, statistical semantics may be used for constructing and examining stimuli and manipulation texts.

## The Potential of Statistical Semantics

### **Objective, systematic and quantitative qualities**

Text data is often thought of as qualitative data. However, in our view, it is not the data per se that is qualitative – what is important is how it is analyzed. Texts may and should to a greater extent be analyzed using quantitative methods, which thus makes them quantitative in nature. Throughout the fields of science, quantification is crucial for making advancements and we argue that statistical semantics allows us to quantify text data, which in turn provides an opportunity for psychological science to leap forward. As a research tool, we have argued that statistical semantics and related methods possess objective, systematic and quantitative qualities. Objective since the analytic procedures limit the subjective interpretations of researchers and allow for independent reproductions. Systematic since all text data methodologically undergoes the same nuanced analytic process using semantic representations covering a large portion of an entire language. Quantitative since statistical semantics is able to model complex interrelationship among words and ultimately enables making statistical inferences. In addition, statistical semantics

methods are data-driven by automatically acquiring this information, whereas predefined word frequency categories require human raters. We believe that these qualities have the potential of rendering statistical semantics an important research tool within psychology and related fields.

### **Validity and reliability**

Numerous research findings based on statistical semantics support the validity and reliability of its application as a research tool. Analyses and methods based on statistical semantics have for instance analyzed texts in order to: predict physical health outcomes such as linking individual narratives to health behaviors (Campbell & Pennebaker, 2003) and linking twitter messages to county level heart disease mortality (Eichstaedt et al., 2015); categorize facial expressions with a significantly higher level of accuracy compared to traditional rating scales (Kjell et al., in revision); predict rating scales of well-being and mental health problems (Kjell et al., in revision); contrast descriptions of psychological constructs with a high level of agreement in relation to prior theoretical understandings (Kjell et al., 2016) and amplify experimental effects (Garcia & Sikström, 2013). Even though some of these studies were exploratory in nature, that statistical semantics has been used to reveal these relationships between text and theoretically relevant outcomes supports its validity.

Results based on statistical semantics also tend to support different kinds of reliability. Kjell et al. (in revision) show that their semantic measures for categorizing facial expressions yield significantly higher interrater reliability compared to corresponding rating scale measures, but also that their semantic measures for subjective reports on mental health demonstrate satisfactory test-retest reliability. Further, Park et al. (2015) show that their language-based assessments for predicting rating scale scores for personality traits yield satisfactory test-retest reliability.

### **Flexibility in collecting and analyzing data**

The broad variety of applications of statistical semantics may further highlight its potential as well as its flexibility as a method. We have discussed several different methods for gathering relevant text data, including asking individuals to describe and narrate their state of mind and experiences, gathering naturally occurring data as well as using text data that is generated or is a part of experimental procedures or stimuli. There are also numerous approaches for quantifying/clustering language, such as LSA or LDA, as well as different ways of analyzing these quantifications,



including semantic t-tests, semantic-numeric correlations, semantic predictions, visualization of keyword analyses, etc.

The analyses and the flexibility incorporated in data-collections unlock many possibilities in terms of testing a wide range of research questions (e.g., the possibility to use already occurring data facilitates longitudinal investigations where one may analyze data created years ago). Furthermore, it is possible to carry out a large number of different analyses on collected data. It is, for example, possible to use different semantic spaces specialized in accommodating different purposes and apply different word norms in order to analyze different aspects of the words/texts.

### **Replications and advancements.**

The applications of statistical semantics have not only led to replication of findings, but also to improved methodologies and new forms of information. For example, Kjell et al. (in revision) not only show that semantic measures may replicate the results of rating scales with a high predictive ability, but also extend traditional methods by describing the measured constructs. They point out that when collecting data in order to capture participants' experiences, traditional methods typically limit participants to express themselves only using closed-ended responses, such as numerical rating scales or fixed response alternatives in the form of checkboxes (i.e., forced choice). Further, the values and understanding of a construct found in the researchers themselves may be imposed in the predefined closed-ended response alternatives (Kjell, 2011), where this method also primes participants with terms necessary for defining the alternatives (Kjell et al., in revision). By not allowing participants to express themselves freely, the closed-ended approach also strips away a great deal of essential and descriptive information. Hence, using statistical semantics may enhance the ways in which we gather and analyze self-reported information concerning individual experiences and states of mind.

Furthermore, Park et al. (2015) show how their language-based assessments using naturally occurring social media texts may be used as an alternative to self-report questionnaires as a "fast, valid, and stable personality assessment" (p. 942). They point out that their method includes advantages such as being unobtrusive, that it may be used retroactively without having to rely on the respondents' memories and that it may generate new insights. For example, Schwartz et al. (2013) state that the automatic clustering of LDA topics facilitated the discovery of unanticipated categories, including sports teams, Japanese cartoons, etc. Kern et al. (2014) further point out that the method of revealing which words are related to which rating scales may provide important information regarding what a questionnaire really measures.

## Limitations and challenges

Potential challenges to the validity and reliability of analyses based on statistical semantics include that to date, automated text analyses fall short when it comes to accounting for word order, words with several meanings, irony and sarcasm. Despite these limitations and as discussed above, statistical semantics has performed well on validation tests as well as in various psychological research settings. Further, in contrast to word frequency strategies, statistical semantics is data-driven by not relying on predefined word categories, and thus encompasses a considerably greater (near complete) portion of a given language. Kjell et al. (in revision) include more than 120,000 words, Park et al. (2015) include more than 51,000 linguistic features (including words, phrases and LDA topics), whereas the LIWC15 default dictionary of Pennebaker and et al. (2015) comprises approximately 6,400 words, word stems and emoticons. In addition, Schwartz et al. (2013) even model words that are frequently/purposefully spelled incorrectly (e.g., *sooo*), slang (e.g., *bestie*, *thingy*), emoticons (e.g., *:)*, which represents a smiling face), abbreviations (e.g., *bday*, *btw*) and symbols oftentimes used on social media (e.g., *<3*, which represents a heart). They also model common two- and three-word phrases, such as *best friend*, *love you*, *life is good* and *why do I*. Hence, even though some nuances of the use of language are lost, statistical semantics approaches cover a large portion of a given language.

Since statistical semantics is based on text corpora written by humans (and word frequency analyses are dependent on human judges), there is a risk of reproducing human-like biases, such as with regard to race and gender. However, methods have been developed for identifying how statistical semantics may replicate human-like biases, for example, by using LSA (Gustafsson Sendén et al., 2015) and word embeddings (Caliskan, Bryson, & Narayanan, 2017). Furthermore, Bolukbasi, Chang, Zou, Saligrama, and Kalai (2016) have developed algorithms that “debias” word embeddings. It is thus possible for research using statistical semantics to utilize these methods in order to identify and limit unwanted biases.

Another challenge concerns being able to efficiently apply statistical semantics within one’s research, considering the fact that it involves new ways of performing analyses. Many of the analyses pertaining to LSA are described and possible to carry out in the online point-and-click software solution that goes by the name of *Semantic Excel* (Sikström, Kjell & Kjell, Appendix B). For useful descriptions of methods related to analyzing big data texts from social media, see Schwartz and Ungar (2015) and Kern et al. (2016).

## Conclusion

Understanding and communicating states of mind using words is an essential quality of human nature, which has largely been an untapped source of data in terms of undergoing sophisticated quantitative analyses in psychology. Statistical semantics has the potential of addressing this. In various research settings, there are likely to be different advantages and disadvantages associated with using the existing variety of statistical semantics methods. Whereas the most suitable approach for text analysis will often depend on the research questions and the type of data, we propose that statistical semantics and related methods represent very useful research tools in psychological science.

## References

- Atkins, D. C., Rubin, T. N., Steyvers, M., Doeden, M. A., Baucom, B. R., & Christensen, A. (2012). Topic models: A novel method for modeling couple and family text data. *Journal of family psychology, 26*(5), 816.
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM, 55*(4), 77-84.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research, 3*(Jan), 993-1022.
- Bolukbasi, T., Chang, K.-W., Zou, J. Y., Saligrama, V., & Kalai, A. T. (2016). *Man is to computer programmer as woman is to homemaker? Debiasing word embeddings*. Paper presented at the Advances in neural information processing systems.
- Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings*. Retrieved from
- Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science, 356*(6334), 183-186. doi:10.1126/science.aal4230
- Campbell, R. S., & Pennebaker, J. W. (2003). The secret life of pronouns flexibility in writing style and physical health. *Psychological Science, 14*(1), 60-65.
- Carlson, T. A., Simmons, R. A., Kriegeskorte, N., & Slevc, L. R. (2014). The emergence of semantic meaning in the ventral temporal pathway. *Journal of cognitive neuroscience, 26*(1), 120-131.
- Chen, M., Mao, S., & Liu, Y. (2014). Big Data: A Survey. *Mobile Networks and Applications, 19*(2), 171-209. doi:10.1007/s11036-013-0489-0
- Dougal, S., & Rotello, C. M. (2007). "Remembering" emotional words is based on response bias, not recollection. *Psychonomic Bulletin & Review, 14*(3), 423-429. doi:10.3758/bf03194083
- Eichstaedt, J. C., Schwartz, H. A., Kern, M. L., Park, G., Labarthe, D. R., Merchant, R. M., . . . Sap, M. (2015). Psychological language on Twitter predicts county-level heart disease mortality. *Psychological Science, 26*(2), 159-169.

- Erk, K. (2012). Vector space models of word meaning and phrase meaning: A survey. *Language and Linguistics Compass*, 6(10), 635-653.
- Foltz, P. W., Laham, D., & Landauer, T. K. (1999). *Automated essay scoring: Applications to educational technology*. Paper presented at the proceedings of EdMedia.
- Furnas, G. W., Landauer, T. K., Gomez, L. M., & Dumais, S. T. (1983). Human factors and behavioral science: Statistical semantics: Analysis of the potential performance of key-word information systems. *The Bell System Technical Journal*, 62(6), 1753-1806.
- Gagné, C. L., Spalding, T. L., & Ji, H. (2005). Re-examining evidence for the use of independent relational representations during conceptual combination. *Journal of Memory and Language*, 53(3), 445-455. doi:<http://dx.doi.org/10.1016/j.jml.2005.03.006>
- Garcia, D., Anckarsäter, H., Kjell, O. N., Archer, T., Rosenberg, P., Cloninger, C. R., & Sikström, S. (2015). Agentic, communal, and spiritual traits are related to the semantic representation of written narratives of positive and negative life events. *Psychology of Well-Being*, 5(1), 1.
- Garcia, D., Kjell, O., & Sikström, S. (2015). A collective picture of what makes people happy: words representing social relationships, not money or material things, are recurrent with the word 'Happiness' in online newspapers. *The psychology of social networking. Personal experience in online community*, 1.
- Garcia, D., & Sikström, S. (2013). Quantifying the Semantic Representations of Adolescents' Memories of Positive and Negative Life Events. *Journal of Happiness Studies*, 14(4), 1309-1323. doi:10.1007/s10902-012-9385-8
- Garcia, D., & Sikström, S. (2014). The dark side of Facebook: Semantic representations of status updates predict the Dark Triad of personality. *Personality and Individual Differences*, 67, 92-96.
- Golub, G., & Kahan, W. (1965). Calculating the singular values and pseudo-inverse of a matrix. *Journal of the Society for Industrial and Applied Mathematics, Series B: Numerical Analysis*, 2(2), 205-224.
- Gustafsson Sendén, M., Lindholm, T., & Sikström, S. (2014). Biases in news media as reflected by personal pronouns in evaluative contexts. *Social Psychology*.
- Gustafsson Sendén, M., Sikström, S., & Lindholm, T. (2015). "She" and "He" in news media messages: pronoun use reflects gender biases in semantic contexts. *Sex Roles*, 72(1-2), 40-49.
- Hansson, K., Bååth, R., Löhndorf, S., Sahlén, B., & Sikström, S. (2015). Quantifying Semantic Linguistic Maturity in Children. *Journal of psycholinguistic research*, 1-17.
- Hart, R. P. (2001). Redeveloping DICTION: theoretical considerations. *Progress in communication sciences*, 43-60.
- Howard, M. W., & Kahana, M. J. (2002). When does semantic similarity help episodic retrieval? *Journal of Memory and Language*, 46(1), 85-98.
- Iliev, R., Deghani, M., & Sagi, E. (2014). Automated text analysis in psychology: methods, applications, and future developments. *Language and Cognition*, 7(2), 265-290. doi:10.1017/langcog.2014.30

- Jameel, S., & Lam, W. (2013). *An unsupervised topic segmentation model incorporating word order*. Paper presented at the Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval, Dublin, Ireland.
- Karlsson, K., Sikström, S., & Willander, J. (2013). The Semantic Representation of Event Information Depends on the Cue Modality: An Instance of Meaning-Based Retrieval. *PLoS ONE*, *8*(10), 1-8. doi:10.1371/journal.pone.0073378
- Kern, M. L., Park, G., Eichstaedt, J. C., Schwartz, H. A., Sap, M., Smith, L. K., & Ungar, L. H. (2016). Gaining insights from social media language: Methodologies and challenges.
- Kjell, O. N. E. (2011). Sustainable Well-Being: A Potential Synergy Between Sustainability and Well-Being Research. *Review of General Psychology*, *15*(3), 255-266. doi:10.1037/a0024603
- Kjell, O. N. E., Daukantaitė, D., Hefferon, K., & Sikström, S. (2016). The Harmony in Life Scale Complements the Satisfaction with Life Scale: Expanding the Conceptualization of the Cognitive Component of Subjective Well-Being. *Social Indicators Research*, *126*(2), 893-919. doi:10.1007/s11205-015-0903-z
- Kjell, O. N. E., Kjell, K., Garcia, D., & Sikström, S. (in revision at *Psychological Methods*). Semantic Measures: Using Natural Language Processing to Measure, Differentiate and Describe Psychological Constructs.
- Kosinski, M., Wang, Y., Lakkaraju, H., & Leskovec, J. (2016). Mining big data to extract patterns and predict real-life outcomes. *Psychological Methods*, *21*(4), 493.
- Landauer, T. K. (1999). Latent semantic analysis: A theory of the psychology of language and mind. *Discourse Processes*, *27*(3), 303-310.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*(2), 211-240. doi:10.1037/0033-295x.104.2.211
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An Introduction to Latent Semantic Analysis. *Discourse Processes*, *25*(2-3), 259-284.
- Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion*, *24*(8), 1377-1388.
- Lapesa, G., & Evert, S. (2014). A large scale evaluation of distributional semantic models: Parameters, interactions and model selection. *Transactions of the Association of Computational Linguistics*, *2*(1), 531-545.
- Lazer, D., Pentland, A. S., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D., . . . Gutmann, M. (2009). Life in the network: the coming age of computational social science. *Science (New York, NY)*, *323*(5915), 721.
- Lench, H. C., Flores, S. A., & Bench, S. W. (2011). Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: a meta-analysis of experimental emotion elicitation. *Psychological Bulletin*, *137*(5), 834.
- Markowetz, A., Błaszczewicz, K., Montag, C., Switala, C., & Schlaepfer, T. E. (2014). Psycho-Informatics: Big Data shaping modern psychometrics. *Medical Hypotheses*, *82*(4), 405-411. doi:<http://dx.doi.org/10.1016/j.mehy.2013.11.030>

- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). *Distributed representations of words and phrases and their compositionality*. Paper presented at the Advances in neural information processing systems.
- Park, G., Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Kosinski, M., Stillwell, D. J., . . . Seligman, M. E. (2015). Automatic personality assessment through social media language. *Journal of Personality and Social Psychology, 108*(6), 934.
- Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). The development and psychometric properties of LIWC2015. *UT Faculty/Researcher Works*.
- Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). Linguistic inquiry and word count: LIWC 2001. *Mahway: Lawrence Erlbaum Associates, 71*, 2001.
- Pennebaker, J. W., Mayne, T. J., & Francis, M. E. (1997). Linguistic predictors of adaptive bereavement. *Journal of Personality and Social Psychology, 72*(4), 863.
- Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology, 54*(1), 547.
- Pennebaker, J. W., & Seagal, J. D. (1999). Forming a story: The health benefits of narrative. *Journal of Clinical Psychology, 55*(10), 1243-1254.
- Rapp, R. (2003). *Word sense discovery based on sense descriptor dissimilarity*. Paper presented at the Proceedings of the ninth machine translation summit.
- Sarwar, F., Sikström, S., Allwood, C. M., & Innes-Ker, Å. (2015). Predicting correctness of eyewitness statements using the semantic evaluation method (SEM). *Quality & Quantity, 49*(4), 1735-1745.
- Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., . . . Seligman, M. E. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PLoS ONE, 8*(9), e73791.
- Schwartz, H. A., Sap, M., Kern, M. L., Eichstaedt, J. C., Kapelner, A., Agrawal, M., . . . Stillwell, D. (2016). *Predicting individual well-being through the language of social media*. Paper presented at the Biocomputing 2016: Proceedings of the Pacific Symposium.
- Schwartz, H. A., & Ungar, L. H. (2015). Data-driven content analysis of social media: a systematic overview of automated methods. *The ANNALS of the American Academy of Political and Social Science, 659*(1), 78-94.
- Stone, P. J., & Hunt, E. B. (1963). *A computer approach to content analysis: studies using the general inquirer system*. Paper presented at the Proceedings of the May 21-23, 1963, spring joint computer conference.
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of language and social psychology, 29*(1), 24-54.
- Turney, P. D., & Pantel, P. (2010). From frequency to meaning: Vector space models of semantics. *Journal of artificial intelligence research, 37*(1), 141-188.
- Weaver, W. (1955). Translation. *Machine translation of languages, 14*, 15-23.







