

# An Experimental History of Positive Psychology

Analysis of Constructs and Instruments

Von der Philosophischen Fakultät der Gottfried Wilhelm Leibniz  
Universität Hannover zur Erlangung des Grades Doktor der  
Philosophie (Dr. phil.) genehmigte Dissertation

von Philipp Eichenberger

2024

Referentin

Univ.-Prof. Dr. Uljana Feest

Korreferentin

Univ.-Prof. Dr. Marianne Sommer

Tag der Promotion: 28.11.2023

<http://dx.doi.org/10.15488/17333>

<https://www.repo.uni-hannover.de/handle/123456789/17461>

This research was funded by the Deutsche Forschungsgemeinschaft  
(DFG, German Research Foundation) – Project 254954344/GRK2073.

تقديم به بهاره و مليکا

**Abstract:**

The aim of this thesis is twofold: To give a historical account of the experiments and development of instruments needed for positive psychology to flourish, and to find a framework within which the epistemic status of another discipline can be assessed.

The narrative follows the operationalizations and measurements of happiness research in psychology. This narrative is then compared with the 'problem of nomic measurement', the feedback loop between theory and measurement, of the historical cases of the theory of temperature and the definition of water as H<sub>2</sub>O.

**Tags:** Positive Psychology, Subjective Well-Being, Happiness

**Schlagworte:** Positive Psychologie, Konstrukt, Glück

# Acknowledgments

This project took many twists and turns and would not have been possible without the many people giving directional advice and emotional support.

First, I have to thank my parents who raised me to be curious.

My gratitude goes to Uljana Feest for the discussions, reading my drafts, the meticulous and precise comments and for the vision guiding this project. Further I thank Marianne Sommer for the kickoff stimulus of the happiness project. With every opportunity I was given to present the progress of my work at the Kolloquium für Kulturwissenschaften at University of Lucerne the project was enriched through the helpful comments by her and the audience.

I thank Morgan Thompson, Leon-Philip Schäfer, Mathias Frisch, Dietmar Hübner and Thorsten Wilholt for the honest discussion on a summary of my thesis.

Many thanks go to my office mates, especially Julia, David/Björn, Ina, Stefano, Daria, Birgit, Jonas, Hannah, Rafael, Leonie and Anna. No thesis would have been written without the coffee breaks and ice creams at mensa. Thank you kindly for the discussions, paper suggestions and the patience needed when I tried to explain the connections between the happiness constructs and some instruments. Further thanks go to the Bielefeld-Crew of GRK 2073, the Vlotho pingpong tournaments will not be forgotten.

Last but not least a few special mentions for people that left their mark on this thesis. I'm indebted to Henrike for her patience while going through my drafts, forcing me to kill my darlings and write more consistently. Thank you Rui for the cake and courage, both were very needed at that time. The 24 Grad meetings with Martin were a stable pillar during my thesis. Thank you for the political discussions, listening to my rants about DB being late again and all the covfefe. And finally, my deepest gratitude goes to the amazing Bahar, without your support this thesis would not have been written.

## List of Abbreviations

**ASQ** Attributional Style Questionnaire

**D-T Scale** Delighted-Terrible Scale (A response scale (grapheme) devised by Andrews and Withey (1976).)

**DSM-IV** Diagnostic and Statistical Manual of Mental Disorders 4th ed. (American Psychiatric Association (1994))

**ESM** Experience Sampling Method

**MTMM** Multitrait-Multimethod Matrix

**NA** negative affect (One of the three components of happiness according to Andrews and Withey (1976). For more information on the third factor please see pages 43 and 48.)

**PA** positive affect (One of the three components of happiness according to Andrews and Withey (ibid.).)

**PANAS** Positive and Negative Affect Schedule

**PERMA** positive emotion, engagement, relationships, meaning and achievement (The components of the new well-being theory introduced by Martin E. P. Seligman (2011).)

**QOL** quality-of-life (The construct measured by the Social Indicators Movement.)

**SWB** subjective well-being

**SWLS** Satisfaction With Life Scale

**VIA-IS** Values in Action Inventory of Strengths

# Contents

<b>Introduction</b>	<b>5</b>
<b>1 How to Perform Research on Science</b>	<b>18</b>
1.1 Taking a Closer Look at Experimentation . . . . .	18
1.2 Van Fraassen's Theory-Laden Measurements . . . . .	25
1.3 On Measurement Scales . . . . .	28
1.4 Active Scientific Realism . . . . .	30
<b>2 The Inception of Subjective Well-Being</b>	<b>33</b>
2.1 Social Indicators Research . . . . .	34
2.2 The Studies Leading to the Definition of Subjective Well-Being . . . . .	37
2.3 The Creation of Subjective Well-Being . . . . .	44
2.4 Conclusion . . . . .	49
<b>3 Research After Establishing the Field - Focus On Positive And Negative Affect</b>	<b>51</b>
3.1 Construct Validation . . . . .	52
3.2 Improving on Known Methods by Adding Variables: Frequency and Intensity	55
3.3 Frequency and Intensity of Affect . . . . .	60
3.4 The Personality Structure of Affect . . . . .	61
3.5 Re-Assessing Frequency of Affect . . . . .	63
3.6 The Iterative Process of Studies and Methods . . . . .	64
3.7 Cultural Influence . . . . .	69
3.8 Putting Things Together . . . . .	69
3.9 The State of SWB Research at the Advent of Positive Psychology . . . . .	70
<b>4 The Rise of Positive Psychology</b>	<b>72</b>
4.1 Short Introduction to Seligman's Work and Theories Before 1998 . . . . .	72
4.2 The Founding of Positive Psychology . . . . .	75
4.3 A Definition of Happiness: The Three Lives . . . . .	76
4.4 Conclusion on The Three Happy Lives . . . . .	80
4.5 Flourish . . . . .	81

<b>5</b>	<b>Methodological Discussion and Criticism of Positive Psychology</b>	<b>84</b>
5.1	On Comparability and Scientific Progress . . . . .	86
5.2	On Validation . . . . .	89
5.3	In Defence of Temporary Claims . . . . .	90
5.4	Assessing Positive Psychology . . . . .	91
<b>6</b>	<b>Conclusion</b>	<b>93</b>



# Introduction

According to Aristotle, happiness is the ultimate purpose of human existence (see Aristotle 2018) and therefore it should be of no surprise that not only greek philosophers but also contemporary scientists devote themselves to studying this topic. Some scholars of happiness and the conclusions of their research have made it to relative prominence lately. The economist Sir Richard Layard for example became famous for advising Tony Blair on his labour market reforms (see Layard 1980, 2005), while the German chancellor Angela Merkel considered the experts from economic and psychological happiness-research in order to determine which policies to apply during her presidency of the G-7 states (see Bräuer 2014, p. 33).

This in itself should not be surprising. In the end we would like politicians to listen more often to science, as the recent Corona pandemic was not shy of reminding us. Further, the scientists themselves seemed to be confident enough in their research that they published guidelines on how to use their indicators for public policy (*i.e.* Ed Diener 2007). Needless to say, the promise to provide instruments allowing to measure happiness and offer evidence-driven interventions to boost it has drawn the attention of politicians and economic decision makers alike to the research of happiness.

This 'movement' of happiness-researchers has been covered in several popular books like *Bright-sided*<sup>1</sup> written by Barbara Ehrenreich (2009), *The Happiness Industry* by William Davies (2015), *Manufacturing Happy Citizens* by Edgar Cabanas and Eva Illouz (Cabanas and Illouz 2019) or *What Is This Thing Called Happiness* by Fred Feldman (2010).

While we have certainly seen research and public policy on happiness rise in recent years, the foundations of this research has been in development for a few decades. The literature ranges back further (*i.e.* W. R. Wilson 1960; Hartmann 1934),<sup>2</sup> but most of the 'founding' texts of today's happiness research in psychology and economics were written in the course of the 1960s to 1980s.

A famous example is the "Easterlin paradox", named after the economist Richard Easterlin

---

<sup>1</sup>The title of the German version of the book, *Smile or Die* (Ehrenreich et al. 2010), might give away the author's assessment of the movement's motto.

<sup>2</sup>In contrary to Wilson and Hartmann, most of the older literature on happiness like Barschak (1951), Beckham (1929), Goldings (1954) and Symonds (1937) does not get mentioned in the literature from the 1960s to 1980s to which in turn contemporary literature is still referring to. For that reason I do rarely consider literature before the 1960s. For a good analysis of subjective measures of happiness starting from the 1920s please see (Angner 2005, p. 27 ff).

who discovered that happiness correlates with income, but stops doing so after a certain income has been reached (see Easterlin 1974).

Approaching this topic, I wondered what this happiness *is* that the scholars of happiness are speaking of. This naïve question is not answered by the scientists explicitly of course, further, every discipline involved applies its own tools and therefore drives the discussion in another direction. For example, some research groups have started using big data analyses, mapping emotionally laden words used on Twitter with levels of heart disease (Eichstaedt et al. 2015a) or tracked whether emotions can be transmitted outside of in-person interactions over Facebook (Kramer, Guillory, and Hancock 2014).

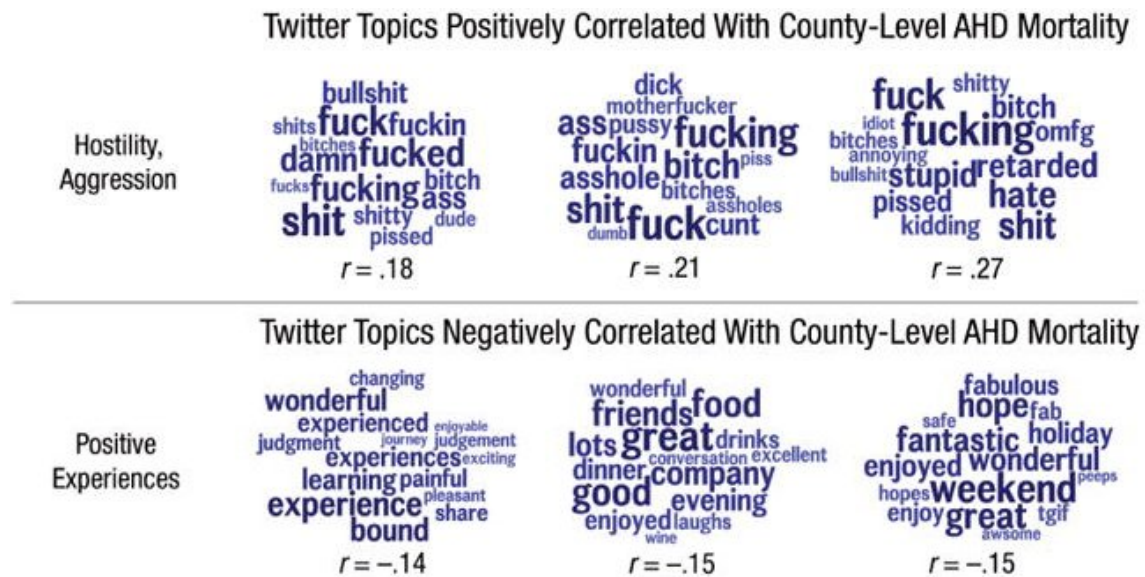


Figure 1: A section of a Twitter wordcloud, correlating emotional word use on Twitter with data on heart disease (Eichstaedt et al. 2015b, p. 158).<sup>3</sup>

While the question what happiness *is* is mostly discussed in philosophy, the main question in science seems to be how to tell that another person is happy. The above mentioned research, while methodologically highly controversial<sup>4</sup> and interesting, still did not come close enough to the third-person assessments of happiness that I was looking for.

Most publications about happiness were written in either philosophy, psychology or economics. While the discussion on happiness in economy certainly revealed interesting results, like its correlation with democracy and federalism (Frey and Stutzer 1999, 2002a), the assessment of happiness is usually done by using a questionnaire, provided by psychologists. Therefore, I have decided to focus on the measurements of happiness, on which the above arguments are based and will for the most part look at literature from psychology and philosophy in this thesis.

<sup>3</sup>Full figure found on <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4433545/figure/F1/>

<sup>4</sup>See the 'Editorial Expression of Concern and Correction of Kramer, Guillory, and Hancock (2014) about the use of Facebook data without informed consent or opting-out options for the users.

The goal of this thesis is to assess happiness as it is being discussed empirically in psychology and stick as close to the psychological literature as possible. That way I aim to lay the experimental mechanics of psychological happiness experiments open. Further, I wish to show how the experimental results changed our perspective on happiness. But most importantly I argue that the understanding of what happiness is or could possibly be had to be changed in order to make happiness measurable and enable experiments in the first place. Lastly, I argue that different measures of happiness should not be confused as measures of happiness but should be seen and referred to as measures of their respective theory or "construct" of happiness. Specific constructs of happiness require a specific viewing angle on happiness that need not necessarily be compatible.<sup>5</sup>

## A Brief Overview and Discussion of the Psychological Literature

In psychology, Mihalyi Csikszentmihalyi coined the term "Flow" in 1975 in his book *Beyond boredom and anxiety: Experiencing flow in work and play* (Csikszentmihalyi 1975) which described a status of total concentration on one activity. Flow has been associated with happiness and Csikszentmihalyi worked on the topic until the end of his career (see Csikszentmihalyi 1990; Csikszentmihalyi and Hunter 2014). The sociologist Ruut Veenhoven published his dissertation on happiness in 1984 (Veenhoven 1984), the same year as the psychologist Ed Diener published his foundation-laying paper "Subjective Well-being" (Ed Diener 1984). Veenhoven is also the founder of the *World Database of Happiness*<sup>6</sup> where all known measures of happiness and studies can be found.

While research on happiness was gaining in popularity, Martin Seligman, a clinical psychologist who has worked on the topics of helplessness and how to strengthen resilience in children (see Martin E. P. Seligman 1975; Martin E. P. Seligman, Reivich, et al. 1995), announced the new movement of "positive psychology" in his presidential address of the *American Psychological Association* (APA) in 1998. This new movement, in contrast to "psychology pre 1998" (Hefferon and Boniwell 2011, p. 5), does not want to mitigate negative effects on the human psyche but reinforce the things that make life worthwhile (see Martin E. P. Seligman 1999a). Seligman's movement soon was joined by Ed Diener and Mihalyi Csikszentmihalyi. The latter co-authored the introduction to the special issue of the *American Psychologist* about positive psychology (see Martin E. P. Seligman and Csikszentmihalyi 2000) and was also involved in writing the *Positive Psychology Manifesto* that was published after the first annual meeting of the founders of this movement (K. Sheldon et al. 2000). In the year 2000, shortly after the establishment of positive psychology, Ruut Veenhoven and Ed Diener co-founded the *Journal of Happiness Studies* together with political scientist Alexandros Michalos (Ruut Veenhoven 2014). Since then, the movement has grown considerably and attracted many new researchers to the field of

---

<sup>5</sup>For a simple example see page 15.

<sup>6</sup><https://worlddatabaseofhappiness.eur.nl/>

happiness-research. Having started in the USA, many universities in the English-speaking world and Europe now host a chair in positive psychology.

In the first few years laypeople probably knew the movement of positive psychology mostly for ranking nations among the happiness of its people. In the year 2000 Ed Diener argued for national indices of happiness providing a current score and helping to track happiness over time (Ed Diener 2000). In 2012 the first World Happiness Index (WHI)<sup>7</sup> was published by the United Nations. A further distinctive attribute is that positive psychology boasts the highest prize money for scientific awards.<sup>8</sup> However, this research of positive psychology is not undisputed.

## Criticism and Contrast to Historical Understandings of Happiness

Critical psychologists and philosophers point to methodological inconsistencies and missing conceptual clarity (Lazarus 2003a,b; Bayertz 2012, 2010; Froh 2004) or even go so far as to call the discipline guilty of positivism, behaviourism and creating a cult of narcissism (see Taylor 2001). Furthermore, anthropologists express concerns about the pressure to show one's own happiness and individuality in modern society, a phenomenon that we could call 'the dictate of happiness'<sup>9</sup> (see Leimgruber 2010).

Who and by what standard one can be considered happy, has changed tremendously in the course of history to which Darrin McMahan's *Happiness - a History* is a true testament (McMahon 2006). The contrast between the way in which happiness has been understood in ancient times to the contemporary understanding of happiness is probably best portrayed by the tale of king Croesus which was written down by Herodotus. The king, convinced that he does not lack anything, asked the wise Solon if he is the happiest person on earth. To his surprise the sage answered: "The thing that you asked me I cannot say of you yet, until I hear that you have brought your life to an end well" (*ibid.*, p. 4). Happiness, in the eyes of Solon the wise, was not something one had worked for in life, but rather the honour and glory that had been bestowed upon a person by the gods at the time of death. Happiness was not considered an emotion but a feature of life as a whole. The state of one person's happiness could therefore not be determined before death. A good death was the best that one could hope for, since afterwards nobody could take the happiness away (*ibid.*, p. 6). Still, happiness was something everyone sought after, but only a small number of 'happy few' were able to reach this half-god status. Only since the time of the enlightenment did happiness come to be understood as something that could be achieved before the beyond (*ibid.*, p. 13). For our modern societies however, the findings of happiness research are of great interest as "all of the important ideologies of modernity are [...] also

---

<sup>7</sup><https://worldhappiness.report>

<sup>8</sup><https://www.apa.org/news/press/releases/2002/05/templeton>

<sup>9</sup>The loosely translated German title of Edgar Cabanas and Eva Illouz book *Manufacturing Happy Citizens, Das Glücksdiktat* (Cabanas and Illouz 2019, 2021).

concepts of happiness” (Leimgruber 2010, p. 48).

It is remarkable that the research of happiness, originally a philosophical-ethical project since Herodotus, has received so much attention in the empirical-experimental sciences. Such a shift in disciplinary interest has of course caught the eye of the social sciences and the humanities. The rise of (self-) therapy since the 1970s has been studied by sociologists and historians alike (for example Maasen et al. 2011; Eitler and Elberfeld 2015; Kleiner and Suter 2015).

Stefanie Kleiner in her work also took a closer look at the humanists in the 20th century who’s topic was picked up by psychologists like Ed Diener and eventually led to positive psychology. She argues the scientific narrative has changed and with it the very essence of what happiness is supposed to be (Kleiner 2015).

## A Brief Introduction in the Recent Philosophical Literature on Happiness

A problem that plagues all disciplines doing research on happiness is the multitudes of terms used for (purportedly) the same thing, *i.e.* happiness and well-being. In philosophy Valerie Tiberius and Daniel Haybron have discussed, within the range of hedonic (subjectivist) and eudaimonic (objectivist) views of well-being, that “[p]ublic decision-making procedures regarding well-being should be subjectivist *in practice*, whether or not well-being really is subjective” (Haybron and Tiberius 2015, p. 714). Both authors have further publications on how a happy life should be lived and how well-being measures should be interpreted philosophically (Haybron 2013, 2010, 2001; Tiberius 2006, 2010; Tiberius and Plakias 2010; Tiberius 2015, 2018).

In his book *What Is This Thing Called Happiness* Fred Feldman argues against the practice of much of the psychologists literature where happiness and well-being (as well as different constructs of the former) are used synonymously. Feldman further argues that there seems to be confusion even among psychologists at the conceptual level amplified by the problem of obscurity (Feldman 2010). He points out that the different terms being used, *i.e.* well-being and happiness, are not interchangeable and articulates his own operationalized and therefore measurable notion of happiness. I will come back to this point on page 12.

Dan Hausman accurately points out, that measurement of well-being are not as clear on what they actually measure as psychologists would like them to be. The question “Overall, how satisfied are you with your life nowadays?”<sup>10</sup> can be answered by referring to one’s momentary feelings, attitude towards one’s life or judgement about one’s life (Hausman 2015, p. 108).<sup>11</sup> ‘Happiness’ therefore potentially means a swath of different things and

---

<sup>10</sup>A question of the “Personal well-being” questionnaire tapping into life-satisfaction. The user guidelines can be found here: <https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/methodologies/personalwellbeingsurveyuserguide>

<sup>11</sup>The difficulty of the terminology also shows in philosophical texts: Hausmann contradicts Haybron and Tiberius separation of well-being and happiness (Haybron and Tiberius 2015, p. 714) since in his chapter “health and happiness” he is referring to subjective well-being several times (Hausman 2015, p. 104 ff).

just by the term alone it is not clear what exactly is being referenced.

Anna Alexandrova in her publications discusses the concepts, measurements and their tests for validation in the empirical research of happiness and well-being. She argues that a contextual concept of well-being might be too narrow from a philosophical point view - but from a practical point of view at least, such a narrow concept renders it measurable (see Alexandrova 2017b, p. 135). Alexandrova further lays out the different possibilities of understanding well-being as a (scientific) object, the different concepts of happiness and what conditions measures of happiness would have to fulfil (see Alexandrova 2012, 2017a, 2018).

While not directly concerned with happiness, Denny Borsboom is criticising psychometric construct validation (Borsboom 2003, 2006; Borsboom et al. 2009), of which happiness-research is certainly a part.

Elina Vessonen is discussing operationalism and validation of tests in psychology and in happiness-research specifically (Vessonen 2019, 2021a).

Further, Erik Angner has written his PhD thesis on the philosophical foundations of subjective measures of well-being and the comparison between them and economic measures for the purpose of public policy (Angner 2005). In further articles Angner argues happiness-researchers differ in notions when they employ the term "subjective well-being" (Angner 2010) and finds them guilty of score reification and literal interpretation fallacy<sup>12</sup> when discussing subjective well-being measures (Angner 2013).

## How to Assess the Psychological Approach to Happiness

At this point it becomes clear that happiness and all the other terms positive psychology has produced cannot be understood from intuition but have to be approached by the way of definition and measurement. We can narrow down the leading question about *what* happiness is to *how* it is being measured and experimented on. In his epistemological reflections on the synthetization of proteins Hans-Jörg Rheinberger argues that no *one* experiment falsifies a theory as Karl Popper has argued (Popper 2008, 2005), rather a whole system of experiments is necessary in order to enable the conditions for new knowledge to appear (Rheinberger 1997). Looking at how experiments function, Rheinberger is interested in how new knowledge can appear, whereas traditional philosophy of science most of the times was merely asking how theories can be tested. The former director of the Max Planck Institute for the History of Science in Berlin has a background in philosophy and later did research on molecular genetics after finishing his thesis and habilitation in molecular biology.<sup>13</sup> With this background Rheinberger argues for a closer look at what is happening materially in a lab when discussing science.

---

<sup>12</sup>In score reification the number resulting from the test is expected to represent the measured attribute perfectly. In trait reification an abstract idea is misinterpreted as a real thing.

<sup>13</sup><https://www.mpg.de/331708/history-of-science-rheinberger>

”If one wants to know how scientific research actually works, one must begin with the characterization of an experimental system, its structure and its dynamics, rather than address theory, or the relation between theory and experiment, as the starting point of the analysis” (Rheinberger 1992b, p. 306).

There of course are differences between the science Rheinberger looked at and the the science of happiness. To begin with, happiness is not a physical material to run experiments on, still, psychologists run labs and perform experiments, inside the lab and 'in the field'. I'd like to take Rheinbergers approach as inspiration to take a closer look at the experiments that psychologists run and the tools they apply to get a better feel for how a certain concept of happiness which is being measured, compared and improved has been conceived.

Hasok Chang also looks at experiments, but does so from a greater distance, both in the historic and in the abstract meaning. In his historical analysis of sciences past, he stresses the operationalizations that were necessary in order to conduct an experiment in the first place (see Chang 2014). A rough idea of the phenomenon alone does not allow the scientist to conduct an experiment, let alone build a measurement instrument. In order for the scientist to succeed, she needs to break down her idea of the phenomenon to concrete manageable pieces that she can manipulate. Chang argues for a kind of coherentism that he calls "epistemic iteration", where scientists might start with an ill-founded guess, but improve on it while their understanding, abilities to manipulate and instruments improve (see Chang 2004, p. 226). What is meant when naming the topic of a scientific endeavour or measurement therefore depends on the operationalization and the current stage of the epistemic iteration.<sup>14</sup> So when we look at statistics showing 'happiness', or any similar term like 'flourishing', we are actually looking at an operationalized term, that simplifies the totality of what 'happiness' could encompass, narrowed down to a specified, single interpretation. The definition of this term is probably well-founded and based on an experimental system defending the claims of the term, differentiating it from other terms.

For example, let's take the list of European countries below, ordered by the percentage of their population that are 'flourishing'. What exactly is being calculated here, such that a holiday destination like Portugal with its abundant sun-hours per day is the least 'flourishing' nation of Europe?<sup>15</sup>

---

<sup>14</sup>For example, today we believe the universe to be cold, except in the places where heat it produced *i.e.* by nuclear fusion in the sun. This was not always the case. Some scientists held that cold was a positive quantity and not just the absence of heat. This dispute has not been settled until the late 18th to beginning of the 19th century (see Chang 2004, pp. 164).

<sup>15</sup>The answer: An operational definition computing the construct scores of competence, emotional stability, engagement, meaning, optimism, positive emotion, positive relationships, resilience, self-esteem and vitality as measured in the European Social Survey (ESS) 2006/2007 (Huppert and So 2013, p. 837, p. 842).

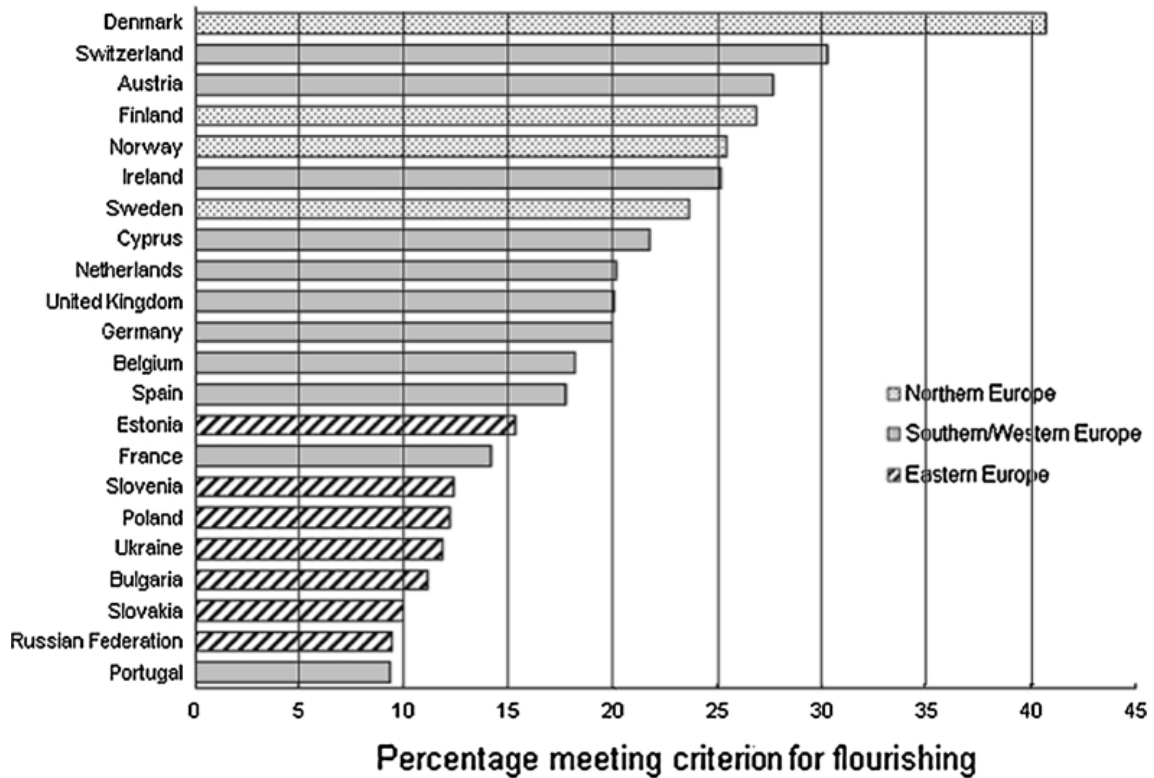


Figure 2: Listing of nations taking part in the European Social Survey according to the operational definition of "flourishing" (Huppert and So 2013, p. 848).

The authors of the study which produced the data seen in this figure do describe how a score of 'flourishing' comes to be, as they made 'flourishing' measurable through a process called "operationalization".

### The Problem With the Terms

Before we discuss *what* and *how* scholars of happiness measure I have to discuss the terminology. A construct (or mental construct) is a label summarizing the underlying cause of covarying behaviours of certain entities under a common term. In physics some well-known constructs are gravity and temperature. The referents of constructs are not measurable directly, however. After observing natural phenomena the commonality of the underlying cause is inferred and summarized under the label of the construct.

"For example, it might be supposed that gravity can be shown by dropping an object to the floor. All that has been demonstrated in that case, however, is the falling of an object, not gravity. Gravity is a label for the hypothetical cause of the falling object, not the observable event" (Binning 2016).

Examples for psychological constructs are intelligence, self-esteem, anxiety and yes, 'happiness'. Being a construct, 'happiness' is not directly measurable. In order for the referents of the construct to be measurable, the construct needs to be operationalized. We could argue



that money has no influence on a person's 'happiness', since 'happiness' is an internal state. We could also argue that having money might allow a person to be more happy, since the person can buy things he needs or likes. However, the construct of 'happiness' is silent on that matter. We could now operationalize 'happiness' as a money-dependent phenomenon and set up an appropriate experiment, or we operationalize it as money-independent and seek empirical evidence for said claim.<sup>16</sup>

There also is a different class of constructs that are partly defined but still lack a distinct operationalization or are only partly operationalized. An illustrative example: Let's define 'true happiness' to be the kind of 'happiness' that is not influenced by money but keep silent on further properties of 'true happiness'. This definition is somehow related to 'happiness', we assume, but slightly more defined. I propose the term "proxy construct" for this kind of more articulated but not completely operationalized construct. We could now operationalize this proxy construct with a very simple instrument and measure 'true happiness' by asking a person "how happy are you?" where the participant scores her answer on a scale from 1 to 3. Such a measure would not exclude money-dependent happiness, but also not obviously tap into it. It is 'one' possible operationalization of 'true happiness', but probably not a very good one. In this example, the proxy construct 'true happiness' would have been operationalized and measured with the "how happy are you" instrument. This instrument consists of only one question (item) that reads "how happy are you" and is scored on a 1 to 3 scale, whereas 3 is the highest score.

In this thesis I will refer to the construct that describes the overarching folk-psychological phenomenon of feeling good and content that lacks a clear way of operationalization as 'happiness'. So, 'happiness' is what you, the reader, intuitively understand when reading that word. Guessing that the reader can neither immediately put said idea into words, not to mention come forward with an operationalization and a measurement, I understand 'happiness' as a vague folk-psychological idea. Whenever I refer to this general, non-measurable idea, it will be marked in scarecrows as '*happiness*'. The totality of research done on any operationalization of 'happiness' is in this thesis referred to as "happiness research" and the scientists conducting the experiments and collecting the data as "happiness scholars". All the psychological constructs covered in this thesis are referring to the folk-psychological idea of 'happiness' to some degree. Happiness in the psychological understanding is a construct that can be operationalized and measured. The happiness construct is described in a theoretical/philosophical way. However, there is another type of constructs as well: These constructs are called "quality of life", "psychological well-being", "subjective well-being", "well-being", "flourishing", "authentic happiness" etc. These constructs are already partly operationalized in the sense that their description include information according to which they must be measured.<sup>17</sup> This information is not

---

<sup>16</sup>See Alexandrova and Haybron (2016, 1103 ff.) for a discussion on the problem of construct validation concerning the money dependent claim.

<sup>17</sup>See page 46 for more information on the conditions that measures of the proxy construct "subjective

exhaustive on how exactly to measure the construct but limits operationalization to some degree. I refer to this type of construct as "proxy construct".

I do realize that my understanding of proxy construct is very similar to Uljana Feests understanding of "concepts". The operationally defined concepts serve as temporary "tools in the generation of knowledge" (Feest 2010, p. 176) and allow the scientists an initial 'take' on their object of interest. The definition of a proxy construct does not need to be well-argued or based on empirical data. As an example, Ed Diener defined that measures of the proxy construct "subjective well-being" have to "cover a period ranging from a few weeks to one's entire life. There is no a priori way to decide what time period is best" (Ed Diener 1984, p. 544). Also, Diener argues for subjective measures and against a third party view where the state of a person is compared against an existing value framework, since this understanding does not line up with "happiness in the modern senses of the word" (*ibid.*, p. 543). While proxy constructs do serve as research tools, I want to stress the argument that proxy constructs refer to a superordinate construct that is less defined. In the discussion about subjective well-being, while not explicitly named, 'happiness' is always the reference.<sup>18</sup>

Another topic worth discussing is how scholars of happiness in psychology and philosophy refer to the respective topic. In the methods section of psychological publications the happiness scholars generally are very clear to what they refer - the measured scores of their instrument, which represents a specific operationalization of a chosen (proxy) construct. The trouble starts when the results are being discussed:

"For simplicity, the rather clumsy term 'reported subjective well-being' will be interchangeably used with the more simple and vivid term 'happiness'" (Frey and Stutzer 1999, p. 756).

"In the literature, terms such as 'happiness', 'subjective wellbeing' [sic!], 'thriving', and 'flourishing' are often used interchangeably, and we use these terms interchangeably here" (Butler and Kern 2016, p. 2).

The conclusions of the studies on an operationalization of 'happiness' very often are applied directly to 'happiness' in the broad sense. This problem of trait reification, the confusion between a measured operationalization of a construct by an instrument and 'happiness', has been criticised by Erik Angner.

"Happiness scholars are guilty of trait reification when they confuse the latent, abstract construct - that which is represented by subjective measures - with happiness in that sense of the word that has moral significance and normative import" (Angner 2013, p. 234).

---

well-being" have to meet as an example.

<sup>18</sup>Even though Ed Diener's research since 1984 almost exclusively revolved around subjective well-being his son named him "the father of modern happiness research" in his obituary <https://positiveacorn.com/blog/ed-diener/>.

Not to differ between a broad construct as 'happiness' and proxy constructs that are in some aspects defined, in my opinion can only produce very imprecise outcomes.

In Philosophy, Valerie Tiberius and Daniel Haybron suggest the following terminology:

"We follow the emerging consensus in using the term 'well-being' for the most general kind of prudential value, or the good for a person. We use 'happiness' in the psychological sense of the term, which could include life satisfaction, domain satisfaction, positive affect, positive emotional condition, etc. We use the terms 'well-being' and 'welfare' interchangeably" (Haybron and Tiberius 2015, p. 714).<sup>19</sup>

The authors choice of terminology is interesting. It suits my intuitive understanding of what "well-being" and "happiness" refer to and is a good fit for a discussion on the difference between a general understanding of "well-being" and psychological "happiness". However, the terminology obscures that *there is no "happiness" in the psychological sense of the term*", but numerous proxy constructs of 'happiness' with potentially contradictory attributes, tailored to the empirical basis and intuitions of the specific research group that is using it.

When deciding to measure 'happiness', happiness scholars first have to decide on a proxy construct, partly limiting possible operationalizations, and then decide on a specific operationalization measured with a specific instrument. Let's look at two different operationalizations of 'happiness': "Life satisfaction" is one of three operationalized factors of the "subjective well-being" proxy construct of 'happiness'. The subjective well-being reading of 'happiness' favors subjective positive whole-life assessments of one's life. One of it's components operationalizations is called "life satisfaction" and is scored by answering five questions, which have to be scored on a scale from 1 to 7. One of the questions of the so-called Satisfaction With Life Scale (SWLS) questionnaire reads as follows:

"If I could live my life over, I would change almost nothing."

Another question reads:

"In most ways my life is close to my ideal." (Ed Diener, Emmons, et al. 1985, p. 72).<sup>20</sup>

This specific operationalization of 'happiness' taps into an individuals cognitive evaluation of her past life as a whole. In contrast to the above, the following operationalization used in a famous study on the heredity of happiness reads as follows:

"Taking the good with the bad, how happy and contented are you on the average now, compared with other people?"

1 = the lowest 5% of the population,

---

<sup>19</sup>A further proposal to note by Feldman (2010, pp. 9–10) is to distinguish between descriptive and evaluative happiness.

<sup>20</sup>The five-item questionnaire "Satisfaction With Life Scale" measuring life satisfaction can be found here: <http://labs.psychology.illinois.edu/~ediener/SWLS.html>

- 2 = the lower 30%,
- 3 = the middle 30%,
- 4 = the upper 30%, and
- 5 = the highest 5%” (Lykken and Tellegen 1996, p. 186).

The scores of the two questionnaires can probably not be compared, since they used different scales, 1-7 and 1-5 respectively. Further, the latter item has participants comparing themselves with others, whereas the questions of the Satisfaction With Life Scale points the participants to their own assessment of their lives. It might be that the phenomenon that these two questionnaires operationalize ('happiness?') overlaps to some degree, but I hope that I have made my point clear that these items potentially lead participants to answer on two only distantly connected questions. This in my view validates the use of two different terms, the respective proxy construct, for what these two questionnaires measure. Subsuming all proxy construct with their operationalizations under one term as Tiberius and Haybron suggest would reduce the accuracy of such a term considerably in my opinion. The proxy construct life satisfaction measured by the Satisfaction With Life Scale refers to a cognitive evaluative reading of one's own life and therefore an 'inward lead' understanding of 'happiness'. The second questionnaire operationalizes 'happiness' comparatively and refers to an understanding of 'happiness' that is relative and dependent on the impression the subjects has of the happiness of her peers. I argue these two understandings of 'happiness' (and hence their operationalizations) are not equal and therefore different terms (*i.e.* the ones of their proxy construct) should be used for them. The empirical research on happiness, even within the discipline of psychology, is a very heterogeneous field. This heterogeneity stems from the vagueness of what 'happiness' could imply and therefore results in vastly different proxy constructs with their own operationalizations and specific instruments used. As Erik Angner has argued, the different accounts of 'happiness' matter, since they are based on differing assumptions about the nature of 'happiness' (Angner 2005, pp. 106–110).

I would have preferred to follow Haybron and Tiberius (2015) suggestion by using "well-being" for the superordinate notion of the good for a person, since that term intuitively describes the topic more accurately. However, this would have brought my use of the terms when describing psychological findings in stark contrast to the terms used in the psychological literature that I am considering for this thesis. The psychologists in their literature are very clear what they want to measure eventually: happiness!<sup>21</sup> Using the term "well-being" for the superordinate notion, while the same term is being used as a more specified proxy construct for 'happiness' by the psychologists would inevitably lead to confusion. Since the aim of this thesis is to analyse operationalizations and proxy constructs of 'happiness' in psychology, the terminology used in this thesis is adapted to

---

<sup>21</sup>In my reading: 'happiness', the overarching folk-psychological phenomenon of feeling good and content that lacks a clear way of operationalization.

the psychological literature cited.

For that reason I will use the term 'happiness' as defined on page 13 marked in scarecrows as the superordinate notion of the good for a person when referring to the general idea without a specific operationalization in mind. In all other instances I will use the name of the specific psychological proxy construct.

## **The Structure of the Thesis**

In the first chapter "How to Perform Research on Science" I will establish my approach, inspired by Rheinberger and others. I argue that taking a closer look at the experiments in happiness research, keeping the different proxy constructs apart and explaining their differences will help understanding the research of happiness better. Further, to my knowledge no publications exists that explains the current state of psychological happiness and happiness research by virtue of experimental history and conclusions drawn. I aim to ameliorate this research gap.

The second chapter, "The Inception of Subjective Well-Being", elaborates the first empirical studies on happiness that paved the way for the contemporary research on this topic. I will argue that these sociological studies including demographic factors laid the groundwork for psychology to be able to access this topic and dominate the discourse by the end of the century.

Chapter three, "Research After Establishing the Field - Focus On Positive And Negative Affect", takes a closer look at the proxy construct of subjective well-being. This proxy construct founded by Ed Diener acted as a loose framework that enabled working on 'happiness' as a psychological topic. I will show by virtue of experiments done and conclusions drawn how 'happiness' became a concept applied to individuals instead of populations, based on personality traits.

The fourth chapter, "The Rise of Positive Psychology", begins with the caesura by the establishment of positive psychology. This marked the big breakthrough of this sub-discipline, accompanied by less empirical groundwork and more focus on interventions.

Chapter five, "Methodological Discussion and Criticism of Positive Psychology", acts as a discussion chapter of the descriptions of experiments, tools and approaches used by the happiness scholars over the decades. Its aim is to provide a basis for further discussions on the quality of the research of happiness and whether *this* happiness is worth pursuing. I hope that this thesis, by providing an accurate historical derivation of the scientific processes that lead to the contemporary proxy constructs of 'happiness' can serve as a basis for further discussions on the epistemic status of 'happiness' as used by happiness scholars.



## Chapter 1

# How to Perform Research on Science

In the introduction I have argued that it is imperative to look at the actual experiments performed and instruments applied when assessing scientific models, theories, or concepts. But are scientific models not well-founded enough? Have scientific theories and models not proved to be accurate by facts and experiments? In this chapter I am going to give a brief introduction about some key approaches of science studies used in this thesis.

### 1.1 Taking a Closer Look at Experimentation

Hans-Jörg Rheinberger argues that there is no such thing like an unfiltered, direct access to the object scientists would like to study. A lot of work has already to be done to enable meaningful scientific experiments. In spirit of the "practical turn" in history and philosophy of science (HPS) Rheinberger rejects a theory-centered analysis of science and argues for a detailed analysis of experimentation when assessing a scientific discipline or topic.<sup>22</sup>

"What goes on *practically* in the research process is the realization, *i.e.* the production, of scientific objects with the help of things that can already be considered and manipulated as sufficiently stable embodiments of concepts, theories, or models (Rheinberger 1992a, pp. 390–393).<sup>23</sup>

Of course, experimentation is not the only legitimate object when assessing science. However, according to Rheinberger, it is experimentation that is the driving force of the research process (see Rheinberger 2018, p. 345). On page 13 I have introduced the high level terminology on the topic of happiness research. Having assessed that experiments will play a vital role in the analysis of the research on happiness I would like to introduce the vocabulary with which to address the experimental components. In order to be able to address the components of experimentation I will rely heavily on Rheinberger's terminology.

---

<sup>22</sup>For an intricate analysis of the practical turn towards (new) experimentalism in HPS and beyond please see Feest and Steinle (2016).

<sup>23</sup>All italics and parantheses in quotes throughout this thesis are by the original authors. Additions to the quotes by the author of this thesis are added in square brackets.

In his book *Toward a history of epistemic things* (Rheinberger 1997) Rheinberger tells the story of the experimental arrangements used in order to do research on cancer cells that eventually ended up enabling to read RNA molecules. Since Rheinberger's vocabulary is used to refer to 'real', physical experiments, it might be a stretch to apply his approach to a research field that is nothing but unphysical.<sup>24</sup> As such, whether and how successful Rheinberger's approach can be applied to 'mental' experiments, is a further research question of this thesis.

## Epistemic Things

Rheinberger introduces a few terms in order to better find our way around in the experimental sciences. The easiest category of entities to describe in Rheinberger's vocabulary are the scientific objects under investigation that Rheinberger aptly names *epistemic things*.<sup>25</sup> Being able to define the "epistemic thing" is the goal scientists set out to achieve. Epistemic things can be of physical nature, a chemical reaction, a psychological mechanism, a biological function. Questions about what exactly the epistemic thing is, are what is driving scientific curiosity. Since the epistemic thing is still under investigation one of its main characteristics is its vagueness, as scientists do not yet know what exactly they are looking for, let alone how to describe it. For this reason, the 'scientific object under investigation' cannot directly be present in the experiments. Would it be possible to describe the epistemic thing in such an accurate manner that it could be included in an experiment, the latter would not be needed any more.<sup>26</sup> In the case of happiness research, 'happiness' is the epistemic thing, which, since it can't be experimented on directly, has to be represented by something else in the experiments. The proxy constructs come in mind, but as I will argue later they represent the same epistemic status as epistemic things and are not measurable either. Therefore, the only thing that can be manipulated and measured in place of the epistemic thing 'happiness' or the epistemic things in the second degree in experiments, the proxy constructs, are the operationalizations of the proxy constructs.

---

<sup>24</sup>Happiness scholars like Martin E. P. Seligman expressed interest in research being able to show happiness as biochemical reactions in the brain. According to the paper "The Neuroscience of Happiness and Pleasure" (Kringelbach and Berridge 2010) neuroscientists have made considerable progress in mapping the neuroanatomy of pleasure. However, I have found very little discussion of neuroscientific discoveries by happiness scholars. Consequently, in this thesis the research on happiness will be treated as a psychological research field only.

<sup>25</sup>In *Toward a History of Epistemic Things: Synthesizing Proteins in the Test Tube* (Rheinberger 1997) the author provides an extensive explication of the different terms and a detailed analysis of the dynamics in science that lead a lab from cancer research to ushering in a new age in microbiology by discovering artificial protein synthesis. The core of Rheinberger's ideas laid out in the book can already be found in a more condensed fashion in some of his previous publications (here, I'll mainly refer to Rheinberger 1992a,b). For this reason I'll preferably refer to these two papers published prior to *Toward a History of Epistemic Things*.

<sup>26</sup>In 1935, the brilliant Ludwik Fleck argued in the same line: "If a research experiment were well defined, it would be altogether unnecessary to perform it. For the experimental arrangements to be well defined, the outcome must be known in advance; otherwise the procedure cannot be limited and purposeful" (Fleck 1979, p. 86).



## Experimental Systems

According to Rheinberger, experimental systems are the smallest functional unit of science.

[The experimental system] is not only a device that generates answers, at the same time, and as a prerequisite, it shapes the questions that are going to be answered. An experimental system is a device to materialize questions” (Rheinberger 1992b, p. 309).

An experimental system, in his view, is more than a mere passive tool that a scientist uses in order to test a hypothesis. It plays an active part in the scientific process. According to Rheinberger, any scientific study actually begins with the choice of such a 'system' (see *ibid.*, p. 312).

The experimental system provides the technical conditions within which an epistemic thing can be framed and operationalized. ”An experimental system creates a space of representation for things that otherwise cannot be grasped as scientific objects” (Rheinberger 1992a, p. 391). A single experiment is only able to give hint at potential answers for a few limited questions. For the most part, every experiment rests on the assumptions and convictions based on prior experiments. Therefore, a single experiment does not show anything, it has to be embedded in a whole system of other experiments in order for it's result to make sense to the scientist.

”Epistemic things take shape in systems of experimentation composed of instruments, apparatus and procedures that stabilize them sufficiently but at the same time allow them to lay out their ambiguity” (Rheinberger 2018, p. 345).

Experimental systems allow to receive gradual hints whether a hypothesis about the nature of an epistemic things should be further pursued or needs to be rejected. It is only within the frame of an experimental system network that an epistemic thing can slowly start to take shape. It is the device for catching a glimpse at the epistemic things that are beyond our knowledge. The experimental systems therefore act as a ”generator of surprises” (see Rheinberger 1992b, p. 307; original argument made in Hoagland 1990, note 9, p. xvii). Rheinberger's experimental system is the antithesis to the conviction that a scientific hypothesis can be rejected or accepted within *one* cornerstone experiment alone.

”In research, *we do not* have to do with single experiments in relation to a theory, but with a whole experimental arrangement designed to produce knowledge that we do not yet have” (Rheinberger 1992b, p. 309).

Many consecutive experiments lead the scientists on a certain trail, convincing them of a specific approach and in its wake potentially lead them to change the discipline in which the research is conducted or bends the research question in a different direction altogether. Only an experimental system, a *system of experiments* in its interlocked totality is able to stabilise the phenomena in order to give convincing answers to scientific questions.

## Technological Objects

Rheinberger argues that one experiment alone cannot prove or disprove a scientific hypothesis. Further he claims that the epistemic thing in a way is not actually part of any experiment, since it yet too vague to be grasped. But how are we supposed to learn something new if the simplest attributes of the epistemic thing are still so unclear? This is where 'technological objects' come into play: These objects are the opposite of epistemic things, they are well-defined and part of the boundary conditions of every experimental system. The behaviour of technological objects should not surprise the scientists, since it is their known characteristics that elevate them to this function. They are the arrangements to refer to when describing the experimental conditions, since the technological objects characteristics are determined. "They perform, at least for the purposes and within the conditions of their use, according to known regularities" (see Rheinberger 1992a, p. 310). The discrimination between epistemic thing and technological object is functional, not material. It is the known characteristics of the technological objects that allow the scientists to describe the observed phenomenon in an experiment and therefore paint the epistemic thing in a less vague manner. Epistemic things ultimately transform into technological objects when "a once surprising result [...] has been sufficiently stabilized" (Rheinberger 1992b, p. 321). After its transformation it can be integrated into a new experimental system as a now stable subsystem.

"Most 'new objects', therefore, are first shaped by 'old tools'. Yet in the long run, a technical system may become completely replaced by systems which embody the current, stabilized knowledge in a more efficient way.[...] As a rule, an experimental system starts as a research device, becomes transformed into a technological device, and finally is replaced" (*ibid.*, p. 323).

It is my understanding that, eventually, devices or instruments are to be categorized as technological objects.<sup>27</sup> As such they are the 'pillars' of every experiment. However, these instruments are not just passively recording data from the experiment, in the contrary. The epistemic things are shaped by the capabilities of the technological object instruments. Rheinberger gives the following example: Philip Siekevitz and Paul Zamecnik from Huntington Laboratories used centrifuges to separate the contents of the cell sap in the process of the establishment of a cell-free protein synthesis system.

"The content of the tissue cells was unpacked into different fractions, which were operationally defined in terms of different centrifugal forces for separating light and heavy components. [...] the centrifugation determined the fractions, and they in turn determined the provisional partition of the scientific object"

---

<sup>27</sup>Rheinberger's technological objects are similar to Bruno Latour's "black boxes" (see Latour 1987). However, the emphasis on technological objects is that all characteristics of interest and the behaviour in the current environment is known. Latour defines a black box as a piece of machinery about which the scientists "need to know nothing but its input and output" (*ibid.*, pp. 2-3).

(Rheinberger 1992b, p. 316).

To put it very bluntly, the adjustment options that the manufacturer of these machines had put on the dial of the instrument in part determined the categorization of the cell contents and played a vital role in which conclusions the scientists were able to draw from these experiments.

## Graphemes

In order to be empirically valuable, the experiment must allow to record some traces of an effect. At some point in the experimental system scientists need to 'see' something they can record: graphemes.

"Here, graphemes are to be understood as the primary, material, significant units of the experimental game which, for the purposes of translation, eventually may become transformed into graphs, i.e. second order representations" (*ibid.*, p. 308).

The epistemic thing is produced or 'seen' in the experimental system thanks to the graphemes which are recorded with the measuring devices built in to the experimental system. The representation of the epistemic thing is constructed from the graphemes. "[T]he representation realized in this space is an epistemic thing which relies on a particular instrument and technical background" (Rheinberger 1992a, p. 393). In some instances in Rheinberger's example experiments, no visible traces could be seen. In these instances markers like colour or radioactive fluorescent materials had to be introduced in the experiments which allowed recordable inscriptions.

"The experimenter is dealing with 'big spots', 'heavy peaks', 'little shoulders', 'yellow soups', and so on. A particular sequence of DNA is 'read' as a ladder of black bars in four adjacent columns on the autoradiograph of a sequencing gel. All measuring techniques and experimental arrangements are optimized in order to reduce noise and to produce traces" (*ibid.*, p. 393).

Since the graphemes are so invaluable to the whole scientific process, not only does the scientist need sophisticated measuring instruments and adds markers where needed, "the scientific object itself is shaped and manipulated with respect to a traceable information" (*ibid.*, p. 393). The scientist needs some kind of instrumental readings in order to push further. In some cases this means the research is going where readings can be produced.

Rheinberger's reflections help to understand how a phenomenon can be produced in such a way that scientists are able to work on it. Questionnaires serve as measuring instruments. Since mental states cannot be directly observed, not unlike many physical phenomena, we need the questionnaire items that, in their totality, might make the phenomenon in question (empirically) visible by providing data about the phenomenon. The actual recording of an effect, in Rheinberger's terms the grapheme, happens on the level of the scales used in

the questionnaire where the study participants jot down the answers to the questionnaire items. The data points collected from the scales are being computed into a questionnaire score which is used for second order representations like graphs and figures.

With Hans-Jörg Rheinberger I have argued that a closer look at the experiments of a scientific endeavour is necessary if we want to know how this science "actually works" (Rheinberger 1992b, p. 306). Rheinberger's literature cited here mostly deals with laboratory experiments. But how does 'nature' end up in the lab in the first place? Science starts before a 'piece of nature' lies on a lab bench.

"Nature as such is not a reference point for the experiment; it is even a *danger*. [...] It is a constant threat of intrusion. [...] If one works with an *in vitro system* [...] one must not contaminate an *in vitro* experiment with 'nature'. So we end up with the paradox that something which we are used to calling 'natural', by remaining so, turns into something artificial. Consequently, the reference point of any controlled system can be nothing else but another controlled system" (*ibid.*, pp. 392–393).

While this section focused on the importance of experiments, I'd like to take a closer look at scientific work pre-experiment in the next section. With Bruno Latour I'd like to argue that experiments and data collections outside of a lab already rely on manipulations in order to be able to abstract and transport what was learned. In order to be able to set up an experiment, the scientists must already have conceptualized 'nature' in such a way that it can be experimented on.

### Latour's Immutable Mobiles

In "Circulating Reference" (Latour 1999) the anthropologist Latour follows a botanist, a pedologist and geographer as a participatory observer on an expedition to the Amazon. The destination of the field trip: The boa vista forest, a site where savanna meets jungle naturally. The goal of this field trip: To establish whether the savanna is eating up the forest or the jungle is slowly extending into the savanna. Upon arrival, Latour recognizes number tags on plants. "I thought I was deep in the forest, but the implication of this sign [...] is that we are *in a laboratory*, albeit a minimalist one, traced by the grid of coordinates" (*ibid.*, p. 32). The botanist picks a few plants from the squares, but not all of them. Only those that she thinks matter for the scientific question at hand. Similarly, the pedologist collects some clods of soil from predetermined depths in the ground and arranges them in a tableau of grids and rows. This tableau "extracts, classifies, and codes the soil" (*ibid.*, p. 56). Conveniently the tableau also acts as a suitcase, allowing ordering and transportation at the same time. Its double function permits the transportation of the soil of the research site and the protocol at which depths the soil was collected allows the comparison of the sample with samples of other sites of interest. In the final report

of the study, the soil and plant samples are processed into a diagram, summarizing what has been started by sorting the soil in the tableau. The material gathered in the research site has been transformed. When a clot of soil has been tested on its consistency, it can be thrown away after the result has been written down. The collected material undergoes many transformations until the study can be published. All these transformations have in common that they are reversible in the sense that the stages of the transformations are traceable in both directions. For any field trip to be successful what has been gathered and displaced has to be presentable to the people that need to be convinced.

Most often this happens by transforming what is to be brought back to an "immutable mobile" (see Latour 1986). Latour argues that every transformation brings material gathered at a research site closer to an immutable mobile. As immutable mobiles, the collected research material can be combined and compared with research material of a different site, since the transformations allow the material to become mobile, immutable and modifiable in scale (see *ibid.*, pp. 18). "A thing can remain more durable and be transported farther and more quickly if it continues to undergo transformations at each stage of this long cascade" (Latour 1999, p. 58).

After a long cascade of manipulations the things gathered at the research site, now in the form of immutable mobiles, can be shown as graphs, figures or pictures. Interesting enough, these visualisations are not trying to be 'realistic'. "[S]cientific imagery is never mimetic. If it were, there would be no gain of information between one step [of transformation] and the next" (Latour 2012, p. 159). At the same time, scientific imagery is very influential and important. It allows to make things present that are far away or unable to be seen by conventional means.



Figure 1.1: Introductory illustration from "The More Manipulations the Better" (Latour 2012, p. 158).

Most scientific arguments are far easier to settle with the appropriate figures and graphs at hand. "[N]o scientific discipline exists without first inventing a visual and written language which allows it to break with its confusing past" (Latour 1986, p. 13).

I take Latour's work as a call to pay great attention to the transformations that scientific material goes through over the course of research. In the case of happiness research this is especially interesting, since in psychology what is being displaced in the field is immaterial. Any visibility therefore must come from transformations. Once immutable and mobile, data turns into figures and illustrations.

## 1.2 Van Fraassen's Theory-Laden Measurements

Psychological phenomena can only be represented by the data provided by the measurement instruments. Constructs are an abstract description of the phenomenon in question but only data according to the operationalizations render the phenomenon tangible in an empirical sense. This is of course true for many sciences, but in psychology it is an especially interesting case, since the effect in a literal sense cannot be 'seen', or the measures seldomly focus on anything that could be visually confirmed. Bas van Fraassen argues, that representation does not depend on resemblance. As the above picture in a jokingly manner pointed at: "It seems then that distortion, infidelity, lack of resemblance in some respect, may in general be crucial to the success of a representation" (Van Fraassen 2008, p. 13).

Therefore, he argues, scientific representation has to be classified at least partly as a cultural object, as instrumentation can only represent the phenomenon symbolically relative to a theory (see *ibid.*, p. 29 and p. 94). Most instruments do not simply allow us visual access to a phenomenon. A measurement is a process, a procedure producing a certain outcome. Instruments are designed to provide measurements for answering specific questions. The information provided by the outcome of the measurement is relevant to the answers of these questions.

Van Fraassen also points out that experiments have further functions than just testing theoretical hypotheses. Rather, he argues, theories are written incrementally and are able to fill a blank in a developing theory. Parameters to be measured cannot be identified without a theory supporting the claim that such a parameter is of any importance. Vice versa, "with no theoretical background, very few procedures classify as measurement procedures" (*ibid.*, p. 123). Every measure therefore adds to and is based on theory at the same time. This is a clear contradiction to Bridgman's operationalism who holds that parameters can be identified through measurement procedures alone, while any concept means "nothing more than a set of operations" (Bridgman 1927, p. 5 ff. see Van Fraassen 2008, p. 124). In a narrow reading of operationalism as a tightly defined scientific philosophy, science can

be done without theory.<sup>28</sup> However, Van Fraassen and many others strongly deny that. In the best case, a measure of an already existing instrument can confirm or make sense of the phenomenon in question, but in some way there is always theory involved coordinating the measurement procedure and the parameter measured.<sup>29</sup>

”So a definite identification, a complete definition, of the measured parameters is possible but only through, at the hands of, and relative to the theory offered and finally accepted to account for the stability of the measurement procedure” (Van Fraassen 2008, p. 124).

In the end it is theory connecting measures to convincing theoretical arguments about how the phenomenon can be explained and which scales do show the explanations best. The procedure of measurement and theory influence each other in such a way over a period of time until the one confirms the other and vice versa. Theory and measurement adapt, ‘get better’, until theory reaches a stable point providing a classification for an object to be measured, before such measures are or can be made (see *ibid.*, p. 143). That is to say, a well-developed/adapted theory to a certain type of measurement can reach a point that theory can predict measurement. Still, the argument stands, that measurements are theory-laden. “[A]n observable (measurable parameter) might not have a specific value outside of the context of measurement” (*ibid.*, p. 150) and by extension, outside of theory. This leads me back to my original, rather naïve research question what *this* kind of happiness is that happiness scholars are speaking of. With reference to van Fraassen I now argue, that this question can only be answered within the space of representation created between theory and measurement.

### 1.2.1 What Do Psychological Measurements Actually Represent?

In the following paragraph I am referring to past discussions in the realm of psychology. I am trying to argue for a theory of measurement that would get us closer to the answer of the question what happiness scholars probably mean, when they refer to a score of a measurement instrument.

Very roughly speaking, research in psychology since the 1960s can be divided in two schools of thought: humanism and behaviorism. Humanists argue, that the whole individual has

---

<sup>28</sup>Bridgeman did not advocate such a stance. In a conference in 1953 he even argued that people are taking things too far, creating a dogma of what he simply deemed to be a point of view (see Chang 2004, p. 142). Simply put, Bridgeman was wary when certain concepts like distance were extended to novel circumstances that required a different set of operations. Originally, distance was measured with a meter rod. ”To say that a certain star is  $10^5$  light years distant is actually and conceptually an entire different *kind* of thing from saying that a certain goal post is 100 meters distant” (Bridgman 1927, p. 18). More on operationalism in this thesis, specifically on Hasok Chang’s understanding of it in chapter 5 from page 84 on. See the [Stanford Encyclopedia article](#) by Chang (2021) for a deep dive.

<sup>29</sup>The story has it that when Galileo showed the imperfections of the sun and the moon as arguments against the Copernican hypothesis, his adversaries did not necessarily deny his observations, but were not convinced by the theory of optics, claiming the devil was inside the telescope producing these misleading images.

to be taken into consideration when discussing a certain phenomenon about a person, with special focus on the person's inner feelings and self-image. Famous protagonists of this movements were Carl Rogers and Abraham Maslow, the latter most famously known for his hierarchy of needs (see Maslow 1943).

Behaviorism was introduced by John B. Watson who envisioned psychology to be "a purely objective experimental branch of natural science" (see J. B. Watson 1913, p. 158). In contrast to the mentalist psychological paradigm of the time and the subsequently established humanistic branch of psychology, introspection is not part of the behaviorists toolbox. B. F. Skinner, a fervent advocate of "methodological behaviorism" argues that no assessment can be made about so-called "private-events" that only manifest in the mental sphere of a person. While Skinner in the direction of the humanists argues for the possible usefulness of self-observation or self-knowledge, he questions the nature of what is felt or observed. On the grounds of missing public agreement on their validity he rules private events out of methodological behaviorism (see Skinner 1976, p. 18). Instead, psychology according to the behaviorist creed should only measure publicly observable behavior and use this data as grounds for a scientific argument. However, many actions in human society are not as clear cut as animal experiments where the subject has to push a lever. Skinner himself argues that humans often act indirectly upon the environment. Instead of looking for water source in order to drink one might simply ask someone for a glass of water. For that reason Skinner argues that verbal behavior is behavior.<sup>30</sup>

Explicitly rejecting behaviorism, the scholars of happiness are proponents of cognitivism which holds mental states accessible *i.e.* through self-reports.<sup>31</sup> However, methodologically cognitive psychology is considered behaviorism in a new form: "The mainstream of psychology [...] remains as firmly behavioristic as it was in 1910" (Leahey 1992, p. 316). While Martin Seligman argues that the shift from behaviorism towards cognitivism was a big step for psychology and attributes his own work on a phenomenon called "learned helplessness" to behaviorism's downfall (see Martin E. P. Seligman 2002, p. 43), John A. Mills sees a major connecting thread between the two: "Today's cognitive psychologists are as deeply committed to operationism, and thus to positivism, as the neobehaviorists were" (Mills 1998, p. 191). Further, Georg Mandler, a cognitive psychologist argues:

"Psychology must talk *about* people. *Your* private experience is a theoretical construct to me. I have no direct access to your private experience. I do have direct access to your behavior. In that sense I'm a behaviorist. In that sense,

---

<sup>30</sup>For more on verbal behavior see (Skinner 1976, pp. 98–99), for more on the indirectness of behavior see (Skinner 1957).

<sup>31</sup>Martin Seligman was working on the topic of depression and discovered a phenomenon he called "learned helplessness" when experimenting on dogs with unavoidable electric shocks. He argues it is difficult "to even approach the sort of phenomena" like learned helplessness from a behaviourist point of view and has "found the cognitive theorizing to be more fruitful and to reflect more accurately those processes that we feel to be reflected in behavior" of the dogs in the experiments (see Maier and Martin E. P. Seligman 1976, p. 41). Seligman later became one of the most cited psychologists to date and is one of the founders of 'positive psychology' studying 'happiness'.



*everybody* is a behaviorist today” (Baars 1986, p. 256).

Since verbal or indirect behavior is considered behavior, a questionnaire asking a person about one’s mental states measuring ‘questionnaire answering behavior’ seems to be in line with methodological behaviorism as it is being used in psychology today.<sup>32</sup>

### 1.3 On Measurement Scales

In 1932 the British Association for the Advancement of Science assigned a committee to debate the problem of measurement representing scientific results from the mathematical, physical and psychological sciences. They came up with four different categories of scales, dubbed ‘Stevens scales’ after the director of the Psycho-Acoustic Laboratory in Harvard. Measurements as per their definition are defined as “the assignment of numerals to objects or events according to rules” (Stevens 1946, p. 677). The representation of scientific data on scales is enabled and limited by the isomorphism between the aspects of measured object and the properties we can accredit the resulting measures.

**Nominal Scales:** These scales categorize equality in one attribute to the objects falling under the same category. One common example is the distinction between ‘verbs’ and ‘nouns’. ‘Happiness’ and ‘measurement’ fall under the ‘nouns’ category, while ‘to feel’ and ‘to be’ would be categorized as ‘verbs’. Nominal scales do not allow any hierarchy between the different categories, ‘nouns’ are not higher valued than ‘verbs’. The only permissible statistics of note are number of cases per category, *i.e.* a text containing X ‘verbs’ and Y ‘nouns’.

**Ordinal Scales:** The classic example for ordinal scales is “Mohs scale of hardness” (see Gregersen and Editors of Encyclopaedia 2022) classifying minerals. Another simple example is a scale representing quality of leather. The measurements that these scales represent allow a hierarchy of the categories. Minerals of category 2 are harder than minerals of category 1, without making any claims that the steps between the categories are equal. The scale is not linear. Also, members of one category only share approximate hardness, determined by observing whether a material is scratched by another substance of defined and superior hardness.

**Interval Scales:** In contrary to ordinal scales, almost all statistical transformations are applicable to interval scales. The main difference to the former is that the difference

---

<sup>32</sup>It has been argued that behaviorism is intrinsically linked with logical positivism and behaviorism therefore inheriting all of logical positivism’s flaws should not be accepted as a valid approach anymore. This is too big a topic to explore in this thesis. Skinner himself argues that behaviorism and logical positivism are concerned with different issues (see Skinner 1976, p. 16). Kurt Danziger, a fervent critic of behaviorism, argues that behaviorism was “deeply beholden to the pragmatic variants of positivism current in North America rather than to the logical positivism imported from Europe” (Danziger 1999). As it seems, behaviorists leaned towards logical positivism for pragmatic reasons without really relying on it. For an excellent and detailed discussion on the connection between logical positivism and behaviorism please see L. D. Smith (1986). On the topic of behaviorism I’d further refer to the interesting discussions by Mills (1998), Danziger (1999) and Bakan (1966).

from one category to the other (the interval) is standardized, hence the name. The difference between a '2' and a '3' on the scale is the same as between a '4' and a '5'. The zero-point in these scales is arbitrarily set by convention, as exemplified in the Celsius and Fahrenheit temperature scales. As such, it is meaningless to say that one value is double to another, such a calculation only makes sense relative to the arbitrarily set zero-point that does not represent a 'real' zero.

**Ratio Scales:** A zero is a 'real nothing' in a ratio scale. Weight *i.e.* in kilos is a common example of a ratio scale. Not only are the intervals standardized, the absolute zero point is also an 'actual absolute zero', no negative is imaginable. Ratio scales are most commonly encountered in physics where operations for all four relations have been determined: rank-order, equality when on the same point of the scale, equality of intervals, and equality of ratio. Scores on these scales can be transformed by all types of statistical means like multiplication, division, mean and median can be applied. Stevens mentions the scale of human observers judging loudness ratios of pairs of tones as one of the known ratio scales of psychological magnitude.

In order to be able to use one scale or another, certain conditions have to be determined. Nominal scales can only be used when at least equality (in the attribute in question) has been determined among objects in one category. In order to be able to use ordinal scales, greater or less has been determined and can be read from the scale. Objects on interval scales additionally have had equality of intervals standardized. Ratio scales have an equality of ratios by having an 'absolute zero' at the bottom of the scale.

Scale	Basic Empirical Operations	Mathematical Group Structure	Permissible Statistics (invariantive)
NOMINAL	Determination of equality	<i>Permutation group</i> $x' = f(x)$ $f(x)$ means any one-to-one substitution	Number of cases Mode Contingency correlation
ORDINAL	Determination of greater or less	<i>Isotonic group</i> $x' = f(x)$ $f(x)$ means any monotonic increasing function	Median Percentiles
INTERVAL	Determination of equality of intervals or differences	<i>General linear group</i> $x' = ax + b$	Mean Standard deviation Rank-order correlation Product-moment correlation
RATIO	Determination of equality of ratios	<i>Similarity group</i> $x' = ax$	Coefficient of variation

Figure 1.2: Stevens scales (Stevens 1946, p. 678).

Stevens holds that most scales used in psychology are ordinal scales (see Stevens 1946, p. 679). As such means and standard deviations ought not to be used, since more knowledge than relative rank-order is necessary to allow these transformations. Stevens allows that pragmatically speaking these calculations on ordinal scales have lead to fruitful results in

many instances but holds that we should be cautious when applying statistics to ordinal scales and what we conclude from them. He goes on to add that percentile measures on rank-order data (as is typical in ordinal scales) by assuming the interval differences are linear (as expected in interval scales) "is, in all strictness, wholly out of bounds", since the linearity of the intervals "is exactly the property which is open to question" (Stevens 1946, p. 679) in ordinal scales. "Most psychological measurement aspires to create interval scales, and it sometimes succeeds. The problem usually is to devise operations for equalizing the unit of the scales" (*ibid.*, p. 679). Stevens argues that usually the problem with using interval scales for psychological measurement is to devise the operations that show the units of the scales are equal. He does allow some leeway however, and argues intelligence is being usefully assessed "on ordinal scales which try to approximate interval scales" (*ibid.*, p. 679).

In general, when discussing psychological data, the question seems to be whether it is permissible to use an interval scale or whether ordinal scales have to be used since proof of the equality of intervals often cannot be given. This position has recently come under attack by Larroulet Philippi (2021) arguing that the strict reading of the conditions of ordinal and interval scales are the two extremes of a continuum of possibilities regarding intervals differences.<sup>33</sup>

While the discussion about scales is closely related to the problem of measures of mental states, I would like to close this discussion in this thesis with one last quote by Stevens:

"Measurement is never better than the empirical operations by which it is carried out and operations range from bad to good" (Stevens 1946, p. 680).

#### 1.4 Active Scientific Realism

Another very important influence on the approach taken in this thesis were Hasok Changs works on temperature and the chemical status of water (see Chang 2004; and Chang 2014). To retell the history on the measurement of temperature and creation of an appropriate scale in very few sentences: Over time and not without controversies natural philosophers decided on fixed points (i.e. freezing and boiling point of water) after which the instruments to measure temperature by the use of various substances (*i.e.* mercury, clay, air) could be developed in order to establish a procedure to represent the various degrees of heat on a scale between the fixed points. But in order to arrive at this stage, several philosophical problems had to be solved. Or, to put it differently, the natural philosophers had to be able to set up experiments, of which the results drawn allowed for a converging answer to the questions they had. Hasok Chang calls this 'the problem of nomic measurement'. We want to measure quantity X, but X is not directly observable. We infer it from quantity

---

<sup>33</sup>For more on the discussion of the Representational Theory of Measurement (RTM) which provides an in-depth foundation for the scales and some new takes on the subject please see (Suppes and Zinnes 1963; Tal 2013, 2020; Vessonen 2021b).

Y, which is directly observable. Therefore we need some sort of 'law', that expresses X as a function of Y.

"The form of this function  $f$  cannot be discovered or tested empirically because that would involve knowing the values of both Y and X, and X is the unknown variable that we are trying to measure. [...] This circularity is probably the most crippling form of the theory-ladenness of observation" (Chang 2004, pp. 59–60).

In the case of temperature, the natural-philosophers soon agreed that instruments can be built by observing the expansion of different materials. Early on, such instruments were built based on the unjustified assumption of uniform expansion of the used materials. But what to do when two instruments disagree with each other? All too often refuting one instrument for another was not an easy matter, since the reasoning alone that an instrument is supposed to work is based on theory. One way to argue for the validity of an instrument is checking whether the instrument gives the same value under the same circumstance when measured multiple times. The assumption here is that "a real physical quantity [like heat] should have one unique value in a given situation" (*ibid.*, p. 77).

In the end, Chang argues, 'inventing' and measuring temperature hinged on the ability of operationalizing the unobservable, to "link abstract theoretical structures to concrete physical operations. [...] Operationalizing an abstract theory involves operationalizing certain individual concepts occurring in it, so that they can serve as clear and convenient bridges between the abstract and the concrete." (*ibid.*, p. 197).

Based on his work on temperature and H<sub>2</sub>O Chang argues for the analysis of scientific developments as "systems of (scientific) practice" that are made up of "epistemic activities" (Chang 2014, p. 15). Chang claims the chemical revolution (away from a phlogistonian theory) was not a scientifically well-argued event and explains the non-acceptance of Johann Wilmelm Ritter's research on batteries and electrosynthesis (which might have saved the phlogiston theory) had to do more with "style and strategy of communication rather than the substance of his work, just as much as Lavoisier's success owed much to his vigorous and effective campaigning" (see *ibid.*, p. 51 and 127).<sup>34</sup> The argument is that temporal assumptions, even when they are ill founded and turn out to be wrong, can lead to scientific progress. The deciding factor is not that those assumptions need to be good. They need to allow operationalizations and measurements where measurements were not possible before. Successful operationalizations made it possible that water came to be understood as an element, the molecule HO and later H<sub>2</sub>O. As it turned out, most of the 'shifts in understanding' were based on "ultimately unjustified" (*ibid.*, p. 133) assumptions, in the sense that alternative theories were as flawed as the ones they were replacing. The difference is that the new theory allowed an operationalization that improved through an

---

<sup>34</sup>No discussion on the point of non-rationality when changing scientific 'paradigms' is complete without reference to Ludwik Fleck who has made similar arguments as early as 1935 based on his work on the history of medical research (see Fleck 1979, 2011).

iterative process.

Hasok Changs "active scientific realism" (see Chang [2014](#), pp.218) serves as a template to analyse operationalizations taking place in the realm of happiness research.



## Chapter 2

# The Inception of Subjective Well-Being

In this chapter I would like to give an overview on the early research on well-being and the eventual emergence of the proxy construct of subjective well-being (SWB). In the narrative of positive psychology, the development of SWB by Ed Diener in 1984 is portrayed as a major milestone. Therefore, the focus of this chapter lies on explaining what SWB as described in 1984 consisted of and why this legacy is so important to positive psychology. The empirical research on 'happiness' by now is to a large degree driven by a psychological 'inward lead' understanding of the topic. In this chapter I am going to argue that the empirical research on 'happiness' actually started with the conviction that demographic measures would enable objective happiness readings of a population. Over time, incremental research lead some proponents in the field to argue that 'happiness' can be measured by psychometrics alone.

I will argue that Ed Diener's paper "Subjective Well-Being" (Ed Diener 1984) received its status not only because of Diener's own studies and experiments, but because it provided a much-needed summary of terms and conclusions in a rather young field of research. I will argue that SWB was an important step of operationalizing 'happiness', without narrowing the topic down to one operationalization only. The legacy of the paper therefore is not that it added another term to the already existing types of well-being but that it unified all or most of the existing terms in one and therefore out of several studies by different research groups created its own field of research within psychology. This, in my opinion, is the power of using proxy constructs. Proxy constructs are often defined pragmatically in tandem with the measurement instruments at hand in the research field. As argued with Hasok Chang, proxy constructs might in part be based on assumptions which in the end turn out to be not true or very inaccurate. Still, if the assumptions allow operationalizations that were not possible before, an iterative epistemic process will improve the assumption. Consequently, several of these proxy constructs are operationally defined and a discussion becomes necessary about the development of the measuring instruments which lead to the very definitions of these proxy constructs. Interestingly, the researchers in the field all seem to agree that their operationally defined proxy constructs relate in some way or form to the state that we call 'happiness' in everyday language. At least at the beginning of this

research, folk psychology must have played an important role by keeping all these different constructs within the imaginative range of what 'happiness' is supposed to be. Only this silent agreement can explain why, in the face of a plethora of different measurement approaches, researchers might disagree on the quality of the results of a measurement scale, while at the same time accepting its results as a measures of 'happiness', the definition of which is still unknown. The scholars of happiness might prefer one over the other proxy construct and argue that a proxy construct might include some unnecessary factors. They however almost always agree that the proxy construct of questionable 'quality' is a proxy construct of 'happiness'. Consequently, another aim of this chapter is to give some insight in the development of these measuring instruments that were improved through several iterations of studies and how 'happiness', this overarching research goal, changed as the measurement methods became more sophisticated.

William Pavot, a psychologist that has worked closely with Ed Diener since graduate school, described measures of SWB to represent subjective indicators of quality-of-life (see Sirgy et al. 2006, pp. 383-384). Subjective indicators in this context are individual evaluations, usually of your own life or mood, as opposed to 'objective' measures like air pollution or literacy rate. Quality-of-life (QOL) in that sense is not only a predecessor of SWB, but SWB a specific, 'subjective' reading of QOL. QOL is a research field that spans several disciplines from psychology to sociology, economics and the health sciences and was initially called Social Indicators Movement. In order to assess SWB we have to understand the origins of the research on quality-of-life.<sup>35</sup> The aim of this chapter is to uncover the links from previous research in the field to what is to become research on SWB.

## 2.1 Social Indicators Research

The research on well-being is tied closely to the social indicators movement that emerged in the 1960s and that is still a relevant research movement today. Most theories and bodies of knowledge that were considered when conceiving SWB, were developed within the social indicators movement. Research findings concerning SWB within populations, the relation of SWB and wealth, reflections on what scales to use and other practical considerations of SWB measurements were frequently published in the movement's *Social Indicators Research* journal by Diener. The movement has strong ties to sociology, economy, psychology, the health sciences, marketing and industrial/organizational psychology. Kenneth Land and Alex Michalos, who both joined the social indicators movement in the 1970s, give a good description of the history and the contents of the social indicators movement (see Land and Michalos 2018), while Sirgy et al. (2006) give a good summary of the breadth of

---

<sup>35</sup>The literature first mentions the "Social Indicators Movement", later the terms "Social Indicators Research" or "Quality-of-Life Research" became more common. I will use the terms interchangeably. To make matters worse, QOL used as the abbreviation for the quality-of-life movements as well as for the proxy construct of quality-of-life. For that reason I will use the abbreviation for the movement only.



the movement. For the sake of context to the concept of SWB, the converging research interests of sociology and psychology are the most important disciplines to consider and will be focused on in this text.

In the early 1960s the US National Aeronautics and Space Administration (NASA) commissioned a study on the consequence of the space program for the population (see Land and Michalos 2018, p. 837). The participating researchers soon realized that they had no data, framework, or methodology at hand with which such a research question could be answered. The realization of this lack of data, framework and methodology sparked the social indicators movement to life. The book *Social Indicators* (Bauer 1966) edited by Harvard business administration professor Raymond A. Bauer, is the product of the attempted study mentioned above and usually seen as the founding document of the social indicators movement.

In the book, Bauer and his co-authors called for the creation of a set of 'true' social indicators. They concluded that purely economic indicators like the gross domestic product (GDP) were not good enough to detect social change. The aim of the social indicators movement was to do research on quality of life of normal people, which is why it is sometimes called "quality-of-life movement" or QOL in short.<sup>36</sup> Economic indicators, in Bauer's view, were not good indicators of quality-of-life since they merely dealt with the quantity of goods and dollars. Soon, the construct of well-being was chosen as describing the overarching research object of the social indicators movement. The researchers became interested in social indicators seeing that projects like the NASA study cannot be conducted without such a frame of reference. The U.S. government also soon became interested in the project. In the same year as Bauer's publication, in 1966, US president Lyndon B. Johnson directed his Secretary of Health, Education and Welfare to develop social statistics and indicators in order to assist the Bureau of Labor Statistics and the Council of Economic Advisors to measure social conditions and progress (see U.S. Department of Health 1969, p. 114). The finished report, published on the last day of Lyndon B. Johnson's presidency, stressed the importance of the so far lacking data and social indicators. While economic indicators were in wide use at the time and the president and his advisors even were required by statute to produce an economic report on the economic health of the nation, no official social indicators had been gathered or evaluated at that point in time. Indeed, economic indicators had become so important for government officials that rising national income was equated with national well-being. However, that this approximation was not working anymore had become all too obvious in the 1960s. While national income was reported to rise rapidly at the time of the 1969 report, the discontent of many citizens was not to miss as violent unrest roared in the streets. This shed serious doubts on economic

---

<sup>36</sup>As far as I can tell both terms, quality-of-life-movement and social indicators movement have been used as early as 1969 interchangeably (see U.S. Department of Health 1969). While some authors today ascribe QOL and well-being a difference in level of evaluation, I did not find any clear differentiation between those terms in the early literature.

indicators being a good measure of how 'well off' the general population actually was. These circumstances were seen as strong arguments for separate indicators and reports on well-being and economy in order to improve public policymaking. In the report from 1969 social indicators were defined as statistic measures of direct normative interest. Therefore, the number of doctors would not be a social indicator, while a figure on health would be. Social indicators therefore were not only supposed to measure clear-cut operationalizations but also represent broad constructs.

By the standards set by this definition most of the contemporary existing social statistics would not suffice as social indicators. According to Heinz-Herbert Noll, member of the advisory board of the government strategy on well-being in Germany, finding a better measure to portray the development of populations than gross domestic product (GDP) was at the heart of the movement since the beginning that has "primarily aimed at operationalizing and quantifying the concept of wellbeing [*sic*]" (Noll 2017, p. 7). Sweden started a similar research endeavor as the social indicators movement in 1965. The Swedish project was based on the question "[w]hat kind of information is needed to give public discussion on [...] social policy an acceptable base?" (Land and Michalos 2018, p. 839). For that purpose, a "level of living" concept was established, which allowed to measure and compare the living conditions of populations. This living concept was supposed to direct political mechanisms that can be used to influence living conditions through social policy. A unitary measure of welfare as well as the attempt to specify "a model for a good life" (*ibid.*, p. 839) were rejected in favor of separate indicators. The importance of the various indicators for public policy was set to be defined through the political process. The last point specifically was interpreted in such a way that the social indicators movement was not interested in philosophical speculations what "the good life" might entail and in specific interventions of improving an individual's life but to prepare the tools for scientifically informed policy making.

Over time a number of indices have been created. For example, the *Better Life Index*, the *Human Development Index* (HDI), the *Canadian Index of Well-Being* (CIW) and the *World Happiness Report* to name just a few. While the movement according to the dominant narrative had its roots in the USA, it produced branches in many places. Most mentioned projects in the reviewed literature were based in North America and Europe, however. In 1974 the *Social Indicators Research* journal was founded with the aim to publish articles related to the measurement of all aspects of quality-of-life.<sup>37</sup> Social indicators were defined as statistic measures of direct normative interest and therefore included classical datapoints like income and education. But as we will see in the following studies, the researchers soon also became interested in producing statistics of the individuals own assessment of their well-being. In summary, the social indicators movement strives to create the tools and frameworks policymakers would need to detect social change. From the beginning a focus

---

<sup>37</sup><https://www.springer.com/journal/11205>

on statistical data, the measure of quality-of-life and well-being, the aim to replace GDP and an adversity to philosophical theorizing was at its core. Even though some scholars had come to the conclusion that the interest in the social indicators movement had waned in the late 1980s and that the field will fade away, interest in has revived and expanded since the mid 1990s (see Land and Michalos 2018, p. 848). Nowadays, the quality-of-life concept receives more attention than ever before, not least thanks to the interest in psychological measures of individual's assessment of their satisfaction with life. In the social indicators movement own representation, the numerous rankings that exist, like "Best Cities for Business" by *Fortune Magazine* or "Best Cities to Raise a Family" by *Readers Digest*, are parallel developments in the private sector that also follow the notion of quality-of-life.

## 2.2 The Studies Leading to the Definition of Subjective Well-Being

In the research that emerged around and after Bauer's 1966 *Social Indicators*, a host of terms like "well-being", "psychological well-being", "perceived well-being", "quality-of-life" and "happiness" were used. These terms were at times used synonymously. At other times the meaning of these terms were operationally defined in terms of a specific scale. So, either the terms referred to 'happiness' in general, or to a specific proxy construct relying on a specific operationalization of said proxy. For example, the Happiness Scale used in one of the studies discussed in this chapter measures a specific operationalization of 'happiness' (the superordinate term) called happiness (the specific operationalization) on a scale from 1 to 3.<sup>38</sup>

One milestone in quality-of-life research leading towards the definition of SWB were the pilot study by Norman M. Bradburn and David Caplovitz (Bradburn and Caplovitz 1965) and the main study *The Structure of Psychological Well-Being* (Bradburn 1969). In the pilot study in 1965 Norman M. Bradburn, a clinical social psychologist, and David Caplovitz, a sociologist, studied four communities in Illinois noting happiness self-reports, marital status, income, education and many more factors of the daily lives of normal people. In the introduction they explained that the study was conducted because the researchers noted the basic lack of knowledge about the behavior of people leading normal lives. The pilot study was published one year before Baur's *Social Indicators* and shared the criticism that without a proper framework social change cannot be detected.

---

<sup>38</sup>In the spirit of the social indicators movement the superordinate epistemic thing should be either referred to as 'well-being' or 'quality-of-life'. In light of too many terms already I will stay with the definition set in the introduction however and will call the superordinate notion of the good for a person 'happiness'.

5. Taking all things together, how would you say things are these days  
 – would you say you're *very happy, pretty happy, or not too happy?*

(Circle one)

Very happy . . . . 1  
 Pretty happy . . . . 2  
 Not too happy . . . . 3

Figure 2.1: Questionnaire item measuring happiness according to (Bradburn and Caplovitz 1965, p. 178).<sup>39</sup>

I would like to focus on two things in this study: First, the definition of the research construct and second the operational definitions and measurements within the study.

Items from Feelings Check-List
<b>Positive cluster</b>
Pleased about having accomplished something
Proud because someone complimented you on something you had done
Particularly excited or interested in something
On top of the world
Angry at something that usually wouldn't bother you
That you had more things to do than you could get done
That you couldn't do something because you just couldn't get going
<b>Negative cluster</b>
Vaguely uneasy about something without knowing why
So restless you couldn't sit long in a chair
Bored
Very lonely or remote from other people
Depressed or very unhappy

Figure 2.2: Bradburn's items measuring affect (Bradburn and Caplovitz 1965, p. 17), yellow emphasis by me.

In order to refer to their research construct either “happiness” or “well-being” were used interchangeably. The construct was defined by Bradburn as the preponderance of positive affect (PA) over negative affect (NA). Affect describes mood or feeling, happiness by this definition therefore consists in having more experiences inducing positive feelings than negative ones. In order to verify this thesis, measures of positive affect and negative affect had to be carried out and compared. A single-item happiness questionnaire (see figure 2.1) was applied, which was meant to measure total happiness of study participants.

A second questionnaire, called “Feelings Check-List”, was to measure positive and negative affect. According to the hypothesis of the authors, these two questionnaires would have to correlate in an expected manner. Subjects had to answer if they had felt at least once as described on the Feelings Check-List as seen on the left during the past week. The items from the feelings check-list<sup>40</sup> did fall into a positive and a negative clusters, meaning the items correlated with other items from the same cluster and the total items of each cluster correlated negatively with the total items of the other cluster.

The second-order representation (the building of the clusters) of the data collected on the graphemes (the scales used to track the answers of the questionnaire items) in-

<sup>39</sup>An illustrative example of the synonymous use of the terms: While the “Happiness-Scale” was meant to measure “happiness” it was described as the person in question would “rate himself in terms of his subjective feelings of well-being” (Bradburn and Caplovitz 1965, p. 7).

<sup>40</sup>The clusters of the “Feelings Check-List” were separated and called Bradburn Positive Affect Scale and Bradburn Negative Affect Scale respectively as seen in later publications like Andrews and Withey (1976).

deed allowed this interesting observation. Further, the comparison of the data from the Feeling Check-List and the Happiness Scale showed that data from the positive and negative cluster correlated to the happiness measure<sup>41</sup> in the expected manner. Respondents that had checked items on the positive cluster more often tended to consider themselves “very happy”, while respondents reporting negative feelings tended to answer “not too happy” on the happiness scale. However, the authors were surprised to find that single items in one cluster were not correlated negatively with single items in the other cluster in a consistent or strong fashion. Bradburn and Caplovitz concluded, that “negative feelings are produced by certain types of experiences the absence of which result not in positive feelings but merely in the absence of negative feelings” (Bradburn and Caplovitz 1965, p. 24). The results of the study supported their thesis of well-being as the preponderance of positive affect over negative affect. The study had additionally shown, that well-being has to be approached by measures of negative *and* positive affect as done in this study and cannot be inferred by an absence of negative feelings alone.

The findings of the above mentioned pilot study from 1965 were replicated in the main study (Bradburn 1969), which is the one usually cited when referring to the independence of positive and negative affect. The term “well-being” is very seldomly used in this publication and is replaced with “psychological well-being” in order to discriminate it from its potential ethical, theological, political or economical understanding. In accordance with the pilot study from 1965 Bradburn wrote in the main study four years later: “the model specifies that an individual will be high in psychological well-being in the degree to which he has an excess of positive over negative affect” (*ibid.*, p. 69). During the four waves of testing the same people, the response of each individual was stable, and the respondents tended to give the same answer they had given previously. While mostly supporting the claims of the pilot study, the main study of 1969 also corrected some popular beliefs about people’s personal attributes and correlations to well-being. Bradburn for example explicitly noted that although women were frequently pictured as being moody, their responses showed greater stability than men’s while also showing similar results in well-being.

Another cornerstone in the development of the social indicators movement was Andrew and Withey’s book *Social Indicators of Well-Being: Americans’ Perceptions of Life Quality* (Andrews and Withey 1976) in which they developed a methodology for the measurement of indicators of well-being as measures of the quality-of-life. These measurement methods would then be implemented in order to create social indicators which could be used in policymaking. As such the authors explicitly counted the study towards the social indicators movement. Indicators should measure something that is cared about directly and over which the government wants to have control. The social indicator is supposed

---

<sup>41</sup>An early example of a confusing terminology: While the Happiness Scale was introduced as a self-report on well-being, its correlation with the Feelings Check-List was described as a correlation with happiness. Later on, Bradburn specified his object of research as “psychological well-being, or happiness” (Bradburn 1969, p. V).

to be defined as at least a part of the operational definition of the concepts central to describing a social system, which differentiates them “from just any social science variable by making it characteristic of the social system” (Andrews and Withey 1976, pp. 3-4). The authors note that quality-of-life is sometimes used as describing an outsider’s judgement of quality of life, measured in noise pollution, crime rate, income levels (objective indicator) etc., but may also refer to privately evaluated aspects of life (subjective indicator). While some connection between objective and subjective measures were expected, Andrew and Withey decide to start measuring well-being by subjective measures. In line with the social indicators movement Andrew and Withey explicitly state that they had made no attempt in their study to make interferences on the structure of perceptions of well-being, since they did not believe that such a structure exists. As such Andrew and Withey kept silent on the subject how people were supposed to live, but generated statistics about people’s life circumstances and their own assessment of how well they were doing. It was decided that two parallel series of social indicator are to be developed: One indicating how people evaluate their own lives, and the other tracking external conditions relevant to these aspects. The focus of Andrew and Withey’s study was to determine indicators of “perceived well-being” and included measuring the importance and interrelations of life concerns like income, family, local government and much more. For the sake of simplicity only the self-evaluation of well-being will be discussed in this section. The authors were led by the hypothesis that a person’s assessment of life quality both includes a cognitive evaluation and an affective (feelings) evaluation. One of the main differences of this study when compared to Bradburn (1969) is that several different questionnaires, most of which not the authors creations, were used as control to show that their own questionnaires provided better results. Further, Andrews and Withey used the same scales as Bradburn in order to measure NA and PA. However, Bradburn had asked *whether* the respondents had felt in the described way, while Andrews and Withey asked how *frequently* they had done so. The results were identical (Andrews and Withey 1976, p. 87).

Answers to questionnaire items (the questions part) were given on different scales (the grapheme, on which the answers can be tracked), typically ranging from one to three, one to five or one to seven. Andrew and Withey developed their own, so-called Delighted-Terrible Scale (D-T Scale) that could be used for different questionnaire items measuring affect and concluded that it had the highest estimated validity on the criteria of “distribution forms, ease of use, and precision of category meanings” (*ibid.*, p. 178).<sup>42</sup>

---

<sup>42</sup>The D-T Scale according to Andrew and Whithey encompasses only the grapheme part tracking the possible answers. It was an improvement over other graphemes, since the seven-point scale allowed more fine tracked answers than an *i.e.* three-point scale and the descriptions for each possible data-point made clearer what each data-point meant. In later publications the D-T Scale was misunderstood as a full questionnaire including a question on happiness (see for example Ed Diener 1984, p. 546; Ed Diener, Emmons, et al. 1985, p. 72). In this thesis I’m trying to be clear in the text when referring to a questionnaire encompassing questions and answers or only to an answering scale, but since several questionnaires have “scale” in their name some confusion might still persist.

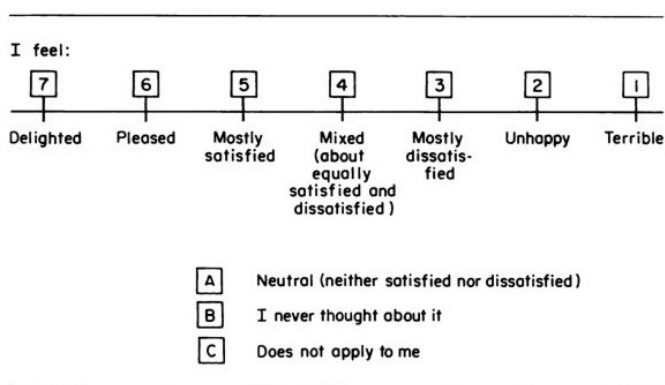


Figure 2.3: The seven-point D-T Scale (Andrews and Withey 1976, p. 18).

tested a multitude of different types of questionnaire items and types of response scales. The D-T Scale represented an attempt to improve over scales used in previous studies, of which most had produced skewed results. For example, studies using the Satisfaction Scale of Campbell, Converse and Rodgers (A. Campbell, Converse, and Rodgers 1976), another research team that worked on similar questions, two thirds of the respondents had checked two of the most happy categories. The satisfaction scale used seven points ranging from “completely satisfied” to “completely dissatisfied” while not having the points in between labelled. Andrew and Withey argued for the superiority of their own D-T Scale over the Satisfaction Scale and others on the grounds that they would expect a lot of people to be more in the middle ground. They felt that the concentration on one end of the scale was due to the other scale’s categories not differentiating between people who were “satisfied” (one of the possible answers) and those who were “merely satisfied” (which was not one of the labelled categories). Their attempt to improve the scale consisted in adding more affect to it by changing the naming scheme of the end categories. Partly convinced by the results of the Satisfaction Scale, the D-T Scale allowed the responses to be made in a range of seven-points in contrary to the many other five and three-point scales. Another consideration that was put into the D-T Scale was the labeling of every possible answering category in contrast to other scales that only labeled the end categories like the satisfaction scale of Campbell, Converse and Rodgers. Andrew and Withey felt this left too much ambiguity as to what the respondents actually meant when checking the middle categories. In order to find valid and efficient ways to measure perceptions of well-being, Andrew and Withey compared several different single-item questionnaires in their study. The national survey was conducted by the Institute for Social Research at the University of Michigan, where both were employed. The data included responses from 8590 individuals and was gathered over the course of six years in 36 states. Additionally, certain use was made from the data of the national survey conducted by Campbell, Converse and Rodgers (*ibid.*).

Of the 37 unique questionnaires used, several were applied multiple times during the two

A comparison between the scales had become necessary, since a multitude of different questionnaires and scales had been used at the time.

The scales compared in this study included the D-T Scale and scales that have been developed and tested by other researchers. These scales were used to measure “quality-of-life”, “well-being” or parts of it. Andrews and Withey

to four hours interviews and some results were calculated into an average of multiple questionnaire responses until a total of 68 measures were conducted. The question of the Satisfaction Scale (A. Campbell, Converse, and Rodgers 1976) read: “how satisfied are you with your life as a whole these days?” and allowed the responses to be made on the above discussed seven-point scale ranging from completely satisfied to completely dissatisfied. Another questionnaire called “Happiness Scale” asked “how do you feel about how happy you are?” and allowed the answers to be made on the D-T scale (Andrews and Withey 1976, p. 67). In these two cases the questionnaires measured satisfaction and happiness, which were operationally defined in terms of the answers on the respective questionnaires. Also included were Bradburn’s PA and NA scales that were discussed above. One scale presented seven faces with varying degrees of smiles and frowns, while asking the question which of these faces comes closest to the respondent’s feelings about his life-as-a-whole. Another scale presented a ladder. The top of the ladder represented the best life the respondent might expect to have, the bottom of the ladder the worst life reasonably to expect. The respondent was asked where she would situate her life most of the time during the past year. A third example presented a “feeling thermometer” with a graphic scale from very cold to negative, very warm and positive at the top. The respondents were asked where they would put their life-as-a-whole on the thermometer. The D-T Scale was also used in order to rate the respondent’s feelings about life-as-a-whole by persons nominated by the respondent and how the interviewers rate the respondent’s feelings about her life as a whole. These examples should suffice in giving a quick glimpse of the breadth of the used scales and questions and the necessity for a comparison of the quality of the data produced by them. The analysis of the results of the questionnaires showed that most of the responses of similar questions tended to cluster together. Interestingly, all the result clusters from less general measures i.e., how the respondent’s life was during the last two weeks were correlated strongest to the cluster of life-as-a-whole responses. This was interpreted as an important finding which showed that all those different questionnaires using a variety of different scales and differently worded questions were “tapping the same underlying phenomenon: general evaluations of life-as-a-whole” (*ibid.*, p. 77).

The cluster of life-as-a-whole responses was named “core cluster” from there on. Andrews and Withey stated that while some of the other questionnaires might be interesting for some aspects of well-being, the questions whose results ended up in the core cluster deserved the primary attention in a social indicators framework (*ibid.*, p. 106). On a technical level, the authors preferred the seven-point over the three-point scales. Both, the Satisfaction and the Happiness Scale questionnaire items were applied with a seven-point and a three-point scales. Correlations were easier to show with data tracked on the seven-point scales. Andrews and Withey argued that three-point scales were too crude, while seven-point scales allowed respondents better to indicate their respective feelings towards the question posed. They also were able to replicate the independence of the positive and negative affect



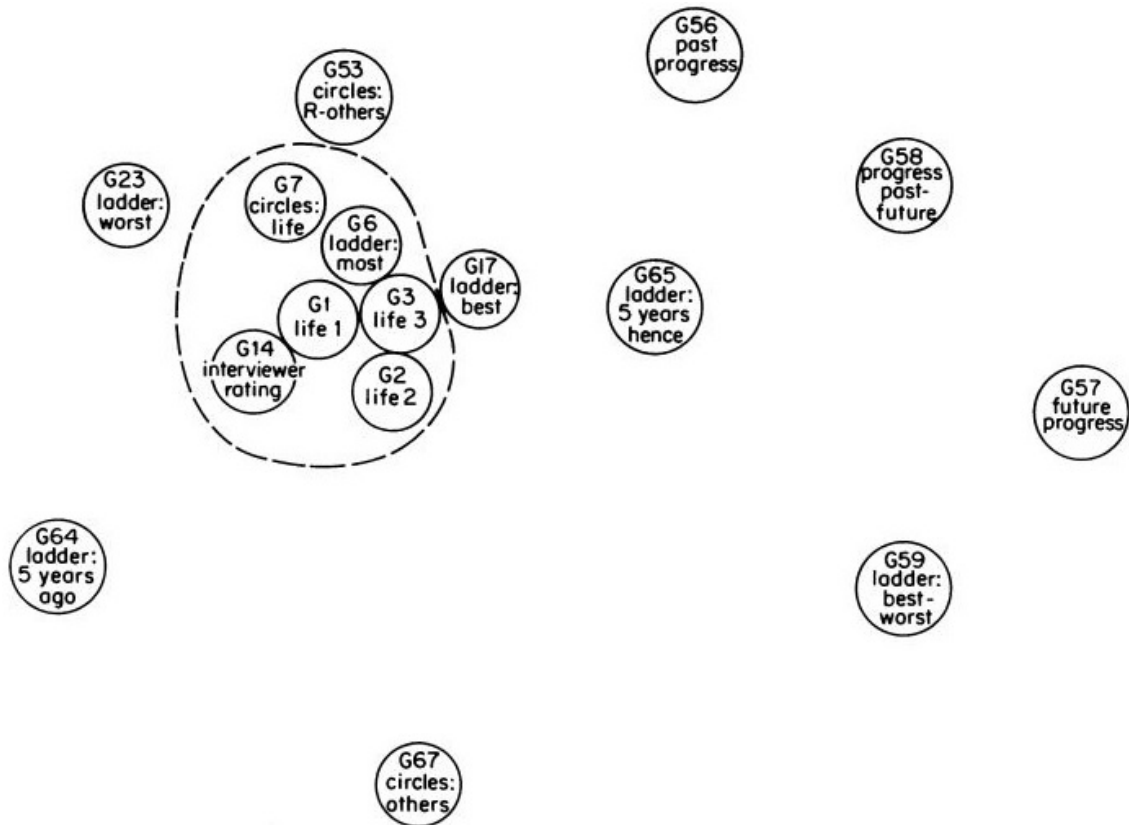


Figure 2.4: The core cluster (encircled) of the life-as-a-whole responses. Each circle with a "G-" title symbolizes the results of a single- or multi-item questionnaire (Andrews and Withey 1976, p. 82).

measures as found by Bradburn. Andrew and Withey noted that the results of the Bradburn PA and NA Scales questionnaires were only moderately related to the results of the general evaluations of life-as-a-whole questionnaires that clumped together in the core cluster and therefore dismissed them as good measures for assessing well-being. This weak correlation opened up a new interpretation however: Since the relationship of affect measures to well-being<sup>43</sup> were very low, this showed that people who reported only few "highs" and "lows" were not worse off in well-being than people reporting a lot of strong feelings on either scale. It appeared that what one feels is more important than how much of it is felt.

Even though Andrew and Withey had concluded that affect was a bad measure for well-being, a conventional factor analysis showed affect to be the second and third factors of assessing life-as-a-whole, while they called the first factor "cognitive evaluation". This first factor was by far more deciding than PA and NA by several orders of magnitude when explaining the statistical variance in life-as-a-whole measures. Andrew and Withey concluded that cognitive evaluation, PA and NA are the three basic components in global measures of well-being.

<sup>43</sup>Here, operationalized by the Happiness Scale and Satisfaction Scale questionnaire results.

The focus of this section has been to demonstrate how the early history of measurement of well-being was rooted in several closely related research programs, which emerged in the context of evidence based social policy. The multitude of scales, questionnaires and terms used reveal the problem researchers face early in the development of a field, where they are trying to develop 'good' measurement instruments, while at the same time not knowing which standards a 'good' measure would have to meet, let alone what exactly to measure. At this stage of a research endeavour, the "epistemic thing", as Rheinberger would call it, is still very opaque. At the same time, Hasok Chang's "problem of nomic measurement" becomes all too obvious, when theory doesn't yet reveal what exactly to measure and existing measures are not well understood enough to argue for a convincing theory yet.

*Social Indicators of Well-Being* by Andrews and Withey was such an important step towards what is to become subjective well-being because it allowed the conclusions drawn on the three factors of happiness by comparing the data of several other questionnaire instruments. On the level of graphemes, *Social Indicators of Well-Being* has also been very influential since the data-quality of the respective graphemes has been compared. To note here is that arguments for dismissing other scales and for creating the Delighted-Terrible Scale was not based on truth or validity. While not questioning the findings presented on the scales showing skewed results, Andrews and Withey "felt this concentration at the »Satisfied« end of the scale posed both conceptual and statistical problems" (Andrews and Withey 1976, p. 19). Hasok Chang claims arguing for "truth" as a leading value when talking about a scientific endeavour is a difficult matter (see Chang 2004, p. 220 ff.). In this vein I find it interesting and telling that parts of an instrument over another were not necessarily chosen because of 'wrong data' or 'the data being not good enough'. Andrew and Withey preferred their D-T Scale not because it was closer to 'reality' than the alternatives. The second order representations based on the data of the D-T Scale grapheme allowed the authors to make additional claims, separating two groups of subjects that were indistinguishable before.

"[W]e suspected that the large group who said they were satisfied included some who were »merely« satisfied and others who experienced a more active and positive enjoyment. We wanted a scale that would separate these groups" (Andrews and Withey 1976, p. 19).

### **2.3 The Creation of Subjective Well-Being**

Before 1984 the term "subjective well-being" had already been used, but the term was not very common. Andrew and Withey claim Campbell, Converse and Rodgers measured subjective well-being with their Satisfaction Scale questionnaire in the book *The Quality of American Life* (A. Campbell, Converse, and Rodgers 1976). However, the authors use the term only twice in the 600 page book and not while talking about their satisfaction scale

questionnaire but in the summary. In most publications, the term “subjective well-being” had been used to stress the subjective nature of certain well-being measures (i.e. R. Larson 1978), while the model of the construct at times seemed to be very similar to Bradburn’s “psychological well-being”, as shown in figure 2.5.

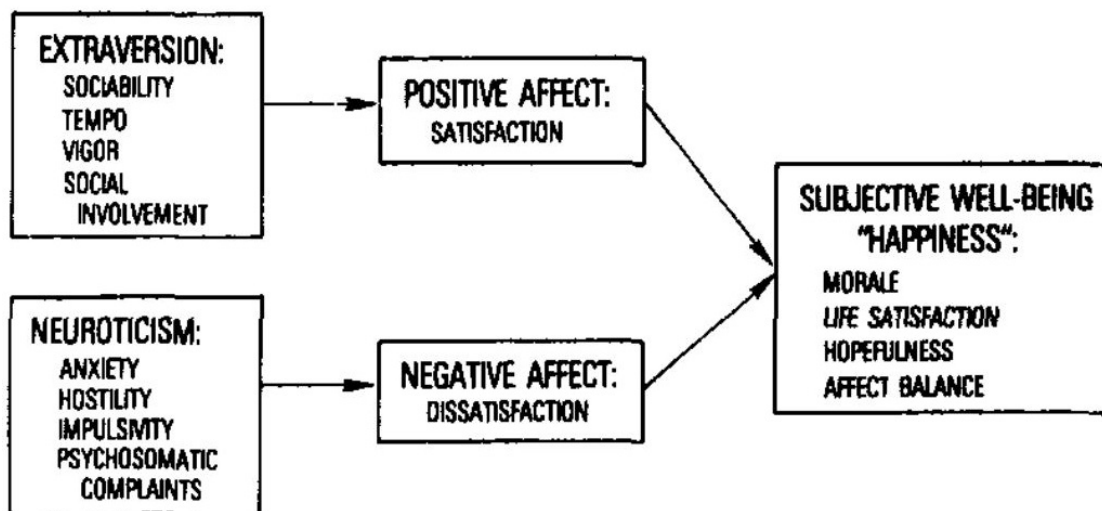


Figure 2.5: The model of subjective-well-being before Diener (P. T. J. Costa and McCrae 1980, p. 675).

The different research groups all seemed to use their own terms but tried to find correlations between the data generated by the different questionnaires. Paul Costa and Robert McCrae refer to Bradburn as well as Andrew and Withey in their study (P. T. J. Costa and McCrae 1980). The term “subjective well-being”, while not exclusively, seems to have been used in gerontology mostly. The above-mentioned Paul Costa and Robert McCrae both worked at the National Institute on Aging at the time of the publication of the paper and Reed Larson also used the term to refer to the happiness of old people. The term “subjective well-being” to that point seems to have been just another term in that area of research that at times has been used as a synonym to ‘happiness’ as was the case for Costa and McCrae.

During the time of the studies by Bradburn Ed Diener was studying psychology at California State University from 1964-1968 and received his PhD at Washington University in 1974. Shortly afterwards Irwin Sarason, one of Diener’s supervisors, together with his co-authors, developed the “life experience survey” (Sarason, J. H. Johnson, and Siegel 1978). This 57 item self-report measure assessed life experiences and their impact on life stress. While no reference to the above-mentioned studies or any of the used terms including “happiness” and “well-being” were made, the study showed somewhat similar results as Bradburn in 1965. Sarason et al. concluded that positive and negative life changes show different patterns of relationship with relevant variables and therefore must be assessed separately. While Sarason and his co-authors might not have been part of

the social indicators movement at the time,<sup>44</sup> their study shows that there had been a general interest in research about how people are coping, which might have affected Diener during his education. At the start of his academic career Diener worked on group dynamics and crowd behaviour, a research field inspired by the riots in the 1960ies and the anti-Vietnam rallies. After receiving tenure at Illinois University in 1980, Diener spent most of his sabbatical year on the Virgin Islands reading into the well-being literature. In his autobiography Diener gives no reason for why he changed the field but that he had been interested in the ingredients of the good life since he read some works of humanistic psychologists and that his parents had taught him to think positively (see Ed Diener 2006). The summary of his sabbatical was published as “Subjective Well-Being” (Ed Diener 1984) which is by now one of the most cited publications in the positive psychology literature.

In this paper in 1984 Diener explicitly defines SWB research to include all constructs described in the literature on happiness, life satisfaction and positive affect. He coherently declares that “a large number of articles devoted to subjective well-being” (*ibid.*, p. 542) have been published in the journal *Social Indicators Research* since 1974, even though few if any of them have used this specific term. Based on Diener’s remarks and William Pavot’s summary of SWB history, proponents of SWB research saw themselves as the psychological branch of the social indicators movement<sup>45</sup> (Sirgy et al. 2006, pp. 383–408). Diener states that literature on SWB is concerned with how and why people experience their lives in positive ways, including cognitive evaluation and affect and so far has used as diverse terms as “happiness”, “satisfaction”, “morale”, and “positive affect”. Concluding from the previous research Diener names three hallmarks of the area of SWB:

1. It is subjective and resides within the experience of the individual.
2. It includes positive measures and does not only consists of the absence of negative factors.
3. Subjective well-being measures include a global<sup>46</sup> assessment of all aspects of a person’s life.

As such, Bradburn (1969) and Andrews and Withey (1976) are explicitly included and named as SWB publications, even though the term was never used in their publications. Furthermore, we see the direct influence these two studies had on the definition on SWB. The second hallmark leans on Bradburn’s purported independence of positive and negative affect, while the third hallmark does justice to Andrew and Withey’s measure of the dominant core cluster of life-as-a-whole measures. Diener also states:

---

<sup>44</sup>The study was published in the *Journal of Consulting and Clinical Psychology* and not in *Social Indicators Research*.

<sup>45</sup>In the mentioned publication *Social Indicators Movement* is explicitly used synonymously with *Quality-of-Life Research Movement*.

<sup>46</sup>Global in this context refers to the total evaluation of one’s life instead of only certain aspects (*i.e.* work life) and implies the assessment of a certain time frame ranging from a few weeks to the whole life (see Ed Diener 1984, p. 544). In that sense it can be compared with the term “life-as-a-whole” in the study by Andrews and Withey (1976).

“[S]ocial scientists have focused on the question of what leads people to evaluate their lives in positive terms. This definition has become labeled life satisfaction and relies on the standards of the respondent to determine what is the good life” (Ed Diener 1984, p. 543).

If a line between the social indicators movement and SWB research needs to be drawn it could be argued that the social indicators movement relies on politics to define 'the good life' using (among others) socio-demographic indicators while SWB research takes its lessons from the social indicators movement but relies on the individual only.

From all the papers and studies about the topic at the time Diener seems to be the first one to discuss several possible understandings of the overarching research object 'happiness' and taps into literature outside of his own discipline. Bradburn for example mentions that “the Greeks” called this subjective sense of well-being “eudaimonia” and argues that he initially expected positive and negative affect to be related based on “general utilitarian notions” (Bradburn 1969, pp. 224–225). The paragraph in which these statements were made deals with the difficulty of naming and measuring the overarching research object without references to back up these claims. Furthermore, philosophers and positive psychologists including Diener would strongly disagree with this interpretation of eudaimonia. Diener in contrast discusses several understandings of the term “happiness” and dismisses Aristoteles virtue-based definition as it does not describe 'happiness' in the modern sense but rather a desirable state judged from a particular value framework. Diener refers to several philosophical publications, some of them being unpublished dissertations. Interesting enough many of these publications are highly critical of the notion that happiness is measurable, referring to the danger of equating pleasure with long term happiness (for example Chekola 1974, p. 19; Michalos 1980).

Outside of philosophy, Diener argues that another social indicators research study by Shin and D. M. Johnson (1978) using the same question for measuring well-being as Bradburn (Bradburn and Caplovitz 1965; Bradburn 1969) established that subjective measures like assessment of needs and perceptions of life standing tend to be more strongly associated with 'happiness' than objective indicators of the resources that the individual has at her disposal. Diener therefore argues that 'happiness' in SWB research is approached either according to Shin and D. M. Johnson (1978) as a global assessment of life quality by self-chosen criteria or as Bradburn's use of the term as a preponderance of positive over negative affect. With reference to Andrew and Withey, Diener names the three components of SWB: (1) life satisfaction, (2) positive affect (PA) and (3) negative affect (NA). While still being interested in affect as his later research shows (*i.e.* Ed Diener and Iran-Nejad 1986; Larsen and Ed Diener 1987), Diener seems to dismiss Bradburn's operationalization of well-being as the preponderance of positive over negative affect and applies Andrew and Withey's discovery of the three components of well-being to SWB. Based on the use of the term by Diener it must be assumed that what Diener calls “life satisfaction” is

what Andrew and Withey had coined “cognitive evaluation”. Like SWB, the term “life satisfaction” was not Diener’s creation either but had already been used in gerontology mostly until that point based on the references given by Diener in “Subjective Well-Being” (see Ed Diener 1984).

After establishing the framework of SWB research, Diener focusses his attention on the operationalization of SWB and compares the already existing questionnaires. As seen in Andrews and Withey (1976) most SWB scales at the time were single-item questionnaires, only consisting of one question. Larsen, Emmons and Diener argued for the inferiority of single-item scales, since they do not allow to assess the separate components of SWB (life satisfaction, positive and negative affect) within that single question. Furthermore, single-item scales are highly affected by the particular wording of the question and most of them so far showed highly skewed results, which the authors didn’t trust (see Larsen, Ed Diener, and Emmons 1985, p. 13). While some multi-item scales had emerged, the items seemed to be tailored to an older population, referring to age and time of life and were therefore not suitable for a general population. Some of the existing multi-item scales also did not line up with Diener’s definition of SWB, since they included factors like zest and apathy (for example Neugarten, Havighurst, and Tobin 1961).

The Satisfaction With Life Scale (SWLS) (Ed Diener, Emmons, et al. 1985) was Diener’s own take at improving the already existing multi-item scales criticised above. The publication was scheduled for 1983 but did not get published until one year after “Subjective Well-Being” (Ed Diener 1984). In Diener’s words this multi-item scale is measuring “global life satisfaction”, which in turn is a component of SWB. Related constructs such as PA are not tapped by the SWLS. According to Diener the SWLS was the first questionnaire that only aimed to tap the judgmental quality (cognitive evaluation) of life satisfaction. The SWLS consists of the following questions:

“Using the 1-7 scale below, indicate your agreement with each item by placing the appropriate number on the line preceding that item. Please be open and honest in your responding. The 7-point scale is: 1 = strongly disagree, 2 = disagree, 3= slightly disagree, 4 = neither agree nor disagree, 5 = slightly agree, 6 = agree, 7 = strongly agree.

- \_\_\_\_ In most ways my life is close to my ideal.
- \_\_\_\_ The conditions of my life are excellent.
- \_\_\_\_ I am satisfied with my life.
- \_\_\_\_ So far I have gotten the important things I want in life.
- \_\_\_\_ If I could live my life over, I would change almost nothing.”

(Ed Diener, Emmons, et al. 1985, p. 72)

From Diener’s description of the SWLS, we can infer that it is only measuring the main component of SWB that Andrew and Withey had called “cognitive evaluation”. It is also

specifically noted, that the SWLS is adequate for all age groups, in contrary to many of the existing multi-item scales of SWB. We further see that the SWLS applies a seven-point scale where all the answer categories have been labelled, similar as the D-T Scale that Andrews and Withey (1976) suggested.

In his study on the SWLS Diener applied several of the questionnaires that he had included in his definition of SWB and compared the results with the SWLS results.<sup>47</sup> The study was to establish the SWLS as a good measure of SWB and compared test results of students and an older population. To Diener's satisfaction the results correlated as expected. This included weak correlations with scales that picked up affect like Bradburn's Positive Affect Scale. This was expected since the SWLS was not intended to pick up much affect since these were the two weak components of SWB.

## 2.4 Conclusion

Diener discussed several theories of well-being and the literature in great length and concluded that so far only a few theories had received propositional development and no attempt to integrate the different theories had been made. He furthermore criticized that constructs should be more rigorously defined as they have been so far in well-being research and that falsifiable propositions must be stated.

I have argued that in his paper "Subjective Well-Being" (Ed Diener 1984) Diener tried to find common ground in a research area riddled with different concepts, scales and repeated name-changes. Andrews and Withey (1976) had accomplished an important step on the way by comparing different types of scales and questionnaire items. While their work certainly was important in terms of 'knowledge on happiness' I would like to stress their work on the graphemes, allowing progress on this end of the experimental systems who's object is to stabilize the epistemic thing 'happiness'. Throughout their arguments were based on a pragmatic spirit, for example arguing for the seven-point D-T Scale, since such data was easier to compare statistically. Further, a long section of their book dealt with reflections on measurement theory, validity and reliability of measures (see *ibid.*, 179 ff. ), a discussion that is seldomly found in comparable literature in the field. Andrews and Withey had proposed good ways of instrument design like using two or three items to assess life-as-a-whole, but did not seem to outline the research area other than through the applied scales and questions.

I argue that the legacy of "Subjective Well-Being" (Ed Diener 1984) consists in allowing a research field to prosper by providing a common ground and a proxy construct to work with, while building on the lessons learned about the core clusters of life-as-a-whole

---

<sup>47</sup>Diener states that he applied Andrew and Withey's D-T scale. Having already established that the D-T scale is only the answering part of the questionnaire it is unclear which question the scale was paired with. From the examples given by Andrews and Withey "how do you feel about your life as a whole?" or "how do you feel about how happy you are?" seem to be the most fitting (for the list of the questions see Andrews and Withey 1976, pp. 66-68).

measurements and instruments designed by Andrews and Withey. The proxy construct of "subjective well-being" was defined broad enough to shelter many different approaches to 'happiness'. At the same time, through its three hallmarks<sup>48</sup> it enforced a psychological reading and operationalization of 'happiness', distinguishing it from the "quality-of-life" field. To that aim, certain terms like life satisfaction and subjective well-being that had mostly been used in specific research areas like gerontology, were broadened to use in well-being research for populations in all kinds of life situations. Diener himself described SWB as an "umbrella term for the different valuations people make regarding their lives" (Ed Diener 2007, p. 152). The SWLS by Diener also implemented many of the findings of previous research endeavours, by including questions concerning global life assessment and focusing on life satisfaction while excluding affect measures in the scale. Summarizing, the definition of SWB respects Andrew and Withey's findings of a "core cluster" when asking people about their life-as-a-whole as opposed to short term life evaluations. While Diener's research will later oppose the alleged independence of positive and negative affect as stated by Bradburn (1969) with reference to weaknesses of the scales used, Diener agrees with Bradburn's interpretation that 'happiness', or in this case the above-defined SWB, cannot be measured by the absence of negative experience alone and therefore must include positive measures (albeit not positive affect). In this sense the outline of SWB presented a Kuhnian paradigm within which a "normal science" could develop (Kuhn 1970). While Diener in his paper regrets the limited theoretical progress made in the field so far, he does not seem to prioritize it either. Diener's outline of SWB is less the outline of a coherent theoretical construct than an outline of how measurements of the most correlated measures of one factor of life-as-a-whole have been achieved so far.

---

<sup>48</sup>Being subjective, including positive measures and including a global assessment of a person's life. For more, see page 46.



## Chapter 3

# Research After Establishing the Field - Focus On Positive And Negative Affect

The aims of this chapter are twofold: 1. To give an overview of how the field of SWB developed. 2. To shed some light on methodological processes undertaken to make sure the respective studies were testing the right hypotheses. Specifically, I will argue that the field of SWB shifted its scientific approach. Subjective well-being, as I described in the last chapter, started as a scientific endeavor based on demographic data converging on a proxy construct based on lessons learned on how the core cluster and how to design instruments. At the end of this chapter, subjective well-being will, at least partly, be based on personality and character traits and therefore conclude on a more 'internal' construct of 'happiness'. The social indicators research movement, the forerunner of the field of subjective well-being, was measuring the influence of demographic factors on well-being by conducting large-scale studies as seen in Andrews and Withey (1976) and Bradburn and Caplovitz (1965) and Bradburn (1969). In these studies, the demographic factors of the population like age, religiosity, education, income relationship status and more were compared to indicators of well-being. Diener's research was part of the social indicators research movement in the sense that some of his papers were published in their journals (for example Balatsky and Ed Diener 1993; Ed Diener 1994, 1995; Ed Diener, Horwitz, and Emmons 1985; Larsen, Ed Diener, and Emmons 1985) and many of his papers included references to publications of the social indicators research movement. The shift from measuring demographic factors to correlating personality factors of happiness was a slow move. First study results were published in the 1960ies, but in the field of subjective well-being the paper of P. T. J. Costa and McCrae (1980) was one of the earliest and most prominent cited that claimed personality trait correlations with subjective well-being. By 1999 Diener described the difference between the conclusions of Warners Wilsons research of the 1960ies (W. R. Wilson 1960; Warner R. Wilson 1967) and his own research as follows:

“Wilson believed the happy person to be well-paid, young, educated, religious, and married. We would emphasise that the happy person is blessed with a

positive temperament, tends to look on the bright side of things, and does not ruminate excessively about bad events, and is living in an economically developed society, has social confidants, and possesses adequate resources for making progress toward valued goals” (Ed Diener, E. M. Suh, Lucas, et al. 1999, p. 295).

In the years after establishing subjective well-being as a field and life satisfaction as its most potent component, it seems to be odd that Diener’s papers after 1984 deal mostly with positive and negative affect, the two minor components of SWB. Those two components of SWB had been defined as the main components of ‘happiness’ by Bradburn (Bradburn and Caplovitz 1965; Bradburn 1969). The third factor had at the time of its discovery by Andrews and Withey not been measured by any instrument yet. But when the relationships of 12 well-being measures that revolved around life-as-a-whole were submitted to factor analysis Andrew and Withey discovered a third factor which they called “cognitive evaluation” (Andrews and Withey 1976, pp. 88–89). This third factor alone contributed to more than 50% of the explained variance, while NA and PA only were about half as strong each. Diener renamed this factor to “life satisfaction” (Emmons and Ed Diener 1985) and I will refer to this factor by this term from now on. Andrews and Withey themselves were very cautious towards life satisfaction. While it was one of the most surprising findings of their study, since this factor had not been discovered before, they mentioned that having a label for this most prominent factor of well-being “may imply a greater conceptual clarity than we believe actually exists” (Andrews and Withey 1976, p. 88). To the author’s knowledge until 1985 no instrument existed to measure life satisfaction/cognitive evaluation until Diener devised the Satisfaction With Life Scale (SWLS) (Ed Diener, Emmons, et al. 1985). With the SWLS, the construct of life satisfaction could be measured directly. All this considered, an explanation is needed why Diener’s team focused on NA, PA and later personality factors like neuroticism and extrovertedness, since his own study results showed life satisfaction to be a more decisive factor. Why were seemingly less-optimal constructs focused?

I hope to show over the course of this chapter that Diener’s primary goal was less to build an intricate theory of happiness and what it might entail, but was research oriented, driven to measure some necessary facts in order to enable building such a theory.

### 3.1 Construct Validation

A “construct” in Psychology is a “theoretical” or “mental” term like intelligence or ‘happiness’. These terms oftentimes have some meaning in folk knowledge, but not necessarily so. They describe phenomena that cannot directly be observed. The measurement therefore needs to be indirect and the construct needs to be operationalized through a test or instrument in order to be measurable. Let’s say a psychologist is doing research on anxiety. She produces an anxiety scale and now she needs to verify whether that scale

actually picks up on the anxiety construct. Since she is teaching at a university and has plenty of students around, exams would make for typical situation where her test subjects are expected to show anxiety. Further, the psychologist expects anxious students at an exam to have sweaty palms. She devises a method to measure how sweaty the palms of the students during an exam were and correlates this with data from the same students applying her anxiety scale just after the exam. Anxiety has now been operationalized by measuring sweaty palms of students during an exam situation and with a self-report anxiety scale. Now the psychologist needs to make sure that both her instruments are actually picking up on the same underlying factor that is anxiety. She compares the data of her two instruments and is happy to find that 81% of the students that said they were anxious during the exam also had sweaty palms. Such a process is called construct validation.<sup>49</sup>

Construct validation is supposed to make sure that the respective instrument actually measures the construct in question. There are different types of validation, but I would like to focus on construct validation here, since this is the only type of validation discussed in the source material. In their paper on construct validation Lee Cronbach and Paul Meehl argued, that such a test should verify that the construct in question is responsible for the variance observed in the test (Cronbach and Meehl 1955). In order to do construct validation a theory is needed that describes the connection between one measured behaviour that is linked to the construct to another observed behaviour also linked to the construct. By checking the expected correlation between two observations, construct validation describes to what extent the construct is responsible for the test score variance. Now one could argue that construct validation does not actually check the proposed link of the construct to the measured attribute. It only checks the correlation of the two measures which the theory suggested is supposed to correlate in a certain manner. This is why construct validation is sometimes referred to as a merely statistical instrument. Construct validation is less of an issue when the construct is operationally defined. In the case of operational definition a mental theory is not operationalized to make the construct measurable, but vice versa. The construct is whatever result a given instrument produces. An example for an operational definition would be: "Intelligence is whatever a given set of IQ measure picks up." Elina Vessonen defends this practice as a temporary workaround. The psychologist "may want to stick with an operationally defined concepts because she does not presently have epistemic access to that »test independent something«" (Vessonen 2019, p. 7).

In their famous study on subjective well-being Andrews and Withey state this variance observed in a test can be distinguished in three parts: a valid component which they want to measure, a correlated error component which includes biases like a tendency of seeing the world positively, and a random error component. The goal therefore is to estimate

---

<sup>49</sup>The above example is a simplified and adapted version of an example given in Cronbach and Meehl (1955, p. 283).

how the total variance can be divided among different measuring methods. “We assume that measures [...] which derive from different measurement approaches, may differ in the proportion of their variance that is valid; i.e., in their »validity«” (Andrews and Withey 1976, p. 179). For that purpose a “validity coefficient”<sup>50</sup> is calculated, which reflects the correlation between the true condition of the construct and the obtained measures. The square of the validity coefficient reflects the proportion of the observed variance that is true variance. This represents the relative amount upon which the test results are measuring the true construct so to speak.

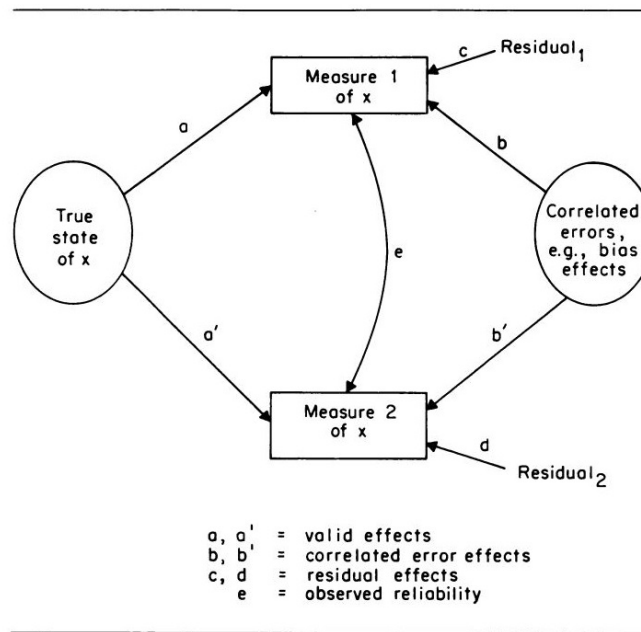


Figure 3.1: "A Rudimentary Measurement Model" by Andrews and Withey (1976, p. 181).

In diagram 3.1 Andrew and Withey illustrate a rudimentary basic measurement model: Two measures of  $x$  are influenced by the same true condition, as well as by some nonvalid effects. Only the link between the two measures  $e$  can be observed, while all other linkages remain undetermined. Andrew and Withey remark, that with some basic assumptions, like there being no method effects and identical validities of the two measures, the validities of the observations can be calculated.

According to Andrew and Withey construct validation requires the estimation of relationships between measures and variables with regards to interrelationships of observed measures and theoretical assumptions about relationships of unobserved variables to the observed ones. Since the publication of Cronbach and Meehl (1955) techniques have been developed to portray theoretical assumptions in structural models and to “compute estimates of construct validity by fitting parameters of the model to the data” (Andrews and Withey 1976, p. 183). According to them, any measurement of a construct can only be

<sup>50</sup>In the social sciences usually Pearsons product moment  $r$  is used.

validated relative to another measure of the same construct.

At the time when Diener started working in the field, NA and PA and their relationship to one another had already been established. That relationship while later refuted as a statistical artifact, had been confirmed by several studies.<sup>51</sup> In stark contrast, life satisfaction was barely picked up by the science community and before Diener devised the SWLS very few publications concerned themselves with that construct. The only other measure that explicitly targeted life satisfaction I came across before 1984 was the "Life Satisfaction Index" by Neugarten, Havighurst, and Tobin (1961), which, according to Diener, is not actually measuring life satisfaction since it is also picking up mood factors and factors like zest versus apathy (see Ed Diener 1984, p. 546). The measure of life satisfaction could not be verified or compared to another measure or data upon which it was not already based on. Therefore, from a measuring perspective, it seemed to be a more fruitful approach to focus on NA and PA in the beginning of SWB research. Measurement instruments like the Bradburn Scales for NA and PA had already been established. Any new developments in measuring instruments could therefore be contrasted to the already existing scales. Furthermore, NA and PA seemed easy enough to tell apart, which helped in validation, while life satisfaction seems a less clear concept to grasp.

### 3.2 Improving on Known Methods by Adding Variables: Frequency and Intensity

Often when researcher criticise a previous study flawed methods are blamed to be responsible for the misleading results. A better method is supposed to produce the results the researcher expects, since the old method is missing a variable that the researcher deems important. In some studies, new variables are added to old methods to prove the point that that this new variable is the missing factor that provoked the criticism. The following study should serve as a twofold example: 1. To show the incremental structure of the psychological studies. 2. to show how adding variables to well-known instruments can lead to different interpretations. In the example of the purported relationship of NA and PA it eventually lead to a different operationalization of the construct of subjective well-being. Bradburn had claimed that NA and PA are not negatively correlated but independent altogether. This phenomenon has been confirmed by the results of many research groups at the time. However, already in his foundational paper "Subjective Well-Being" Diener challenged the alleged independence of PA and NA. According to him, NA and PA, as represented in Bradburn's study, did not paint a full picture. He criticised that the measures based on Bradburn's scales only track the *occurrence* of PA and NA but neither the strength (the *intensity*) nor how often it happened (*frequency*). The paper "The Independence of Positive and Negative Affect" (Ed Diener and Emmons 1985) summarizes

---

<sup>51</sup>More on the purported independence of NA and PA on page 58.

a total of 5 studies. These studies incrementally test Diener's own thesis on how he expected NA and PA to correlate, while at the same time testing how the methodological changes influence data gathered with Bradburn's scales. Every study introduces small changes to the original study procedure of Bradburn and checks whether the change in method produces data that lines up more with Diener's thesis than with Bradburn's.

Study 1 confirmed the independence of NA and PA and improved on Bradburn's existing instruments. Bradburn gave no information upon which basis his scale has been constructed.<sup>52</sup> Bradburn merely mentions that the items "were chosen to represent a wide range of pleasurable and unpleasurable experiences apt to be common in a heterogeneous population" (Bradburn and Caplovitz 1965, p. 16). In order to check whether the specific items in Bradburn's scales were responsible for the results, Diener devised a new scale with emotions thought to be related to 'happiness' and unhappiness. This scale was based on lists of frequent emotions comprised by other research teams who were working in the field of emotion theory (Plutchik 1980; James A. Russell 1979; James A Russell and Mehrabian 1977).<sup>53</sup> The groups in the first study reported the intensity in which they had felt certain emotions during the past year with Diener's new scale and the frequency how often they had felt the emotions with Bradburn's PA and NA scale. The two scales measure frequency and intensity because Diener's scale asks participants on a scale from 1 to 7 how intense they had felt the emotion while Bradburn's scale asks participants to respond how often they had experienced said emotions. The study confirmed the independence of the two factors (NA and PA) in frequency and intensity in the respective scales and showed the factor loadings of the items to the two factors with Diener's scale.<sup>54</sup>

The third factor discovered by Andrew and Withey (cognitive evaluation/life satisfaction) was on purpose excluded in this study since the aim was to test Bradburn's independence of NA and PA with Diener's new scale. According to study 1 a scale which's items were more thoroughly picked and checked on intensity instead of frequency of emotions produced the same results as Bradburn's scale. Study 2 consisted of the almost identical procedure as study 1, but this time the timeframe was changed. The participating students reported the frequency (Bradburn's scale) and intensity (Diener's scale) of their emotions for the past year and the past month. While the intensity results over a year were similar to those of study 1 and showed independence, the results over the timeframe of a month showed that NA and PA indeed did correlate negatively. Also when measuring NA and PA

---

<sup>52</sup>The items on the scale were "pleased", "proud", "excited", "on top of the world", "angry", "too much to do", "couldn't get going", "uneasy", "restless", "bored", "lonely", "depressed" (Bradburn and Caplovitz 1965, p. 17). The emotional terms were presented within a sentence. Participants had to respond whether they had not felt that emotion during the timeframe at all, "once", "several times" or "often".

<sup>53</sup>In Diener's new scale the terms were "happy", "joy", "pleased", "enjoyment/fun", "glad", "delighted", "contented", "angry", "fear/anxiety", "frustrated", "depressed", "annoyed", "sad", "gloomy". Participants had to rate the intensity of their mood on a scale from 1 (not at all) to 7 (very much) (Ed Diener and Emmons 1985, p. 1108). As argued in chapter 1, scales from 1 to 7 were considered superior to other scales and very popular in the field.

<sup>54</sup>"Delighted", for example had a factor loading of 74 PA and -15 NA, whereas "fun" in the same subject sample had a factor loading of 78 PA and 0 NA.

frequency with Bradburn's scales over the timeframe of a month, a negative correlation was observed, albeit with low significance. Study 2 therefore showed two things: Since the student sample consisted of black Caribbean students on the Virgin islands and frequency and intensity results over a year were similar to those of the University of Illinois students in study 1, cultural bias could be expected to be minimal. Second, frequency and intensity results over the shorter timeframe of a month did show negative correlation. However, the results could be influenced by memory distortions since people might recall memories selectively.

Study 3 sought to remedy this bias: Participants reported the intensity of their moods with Diener's new questionnaire during the day and in the evening and had to hand in the questionnaires the next morning in order to ensure daily completion. While the questionnaire items were partly changed based on the factor loadings of the previous studies, what is interesting to us here is the change in methodology in order to reject memory bias. With this method the team not only had access to the intensity of certain emotion of the participants summarized over a day but also how the intensity of the emotions changed on a daily basis. As such the researchers also had access to the frequency of the emotions, without explicitly asking for it. When calculating frequency NA and PA correlations of days within-subject, PA and NA were negatively correlated, confirming study 2. So, when a person feels PA, she is less likely to feel NA at the same time. However, when comparing how much NA and PA people felt overall in average (intensity), the results were positively correlated. People that feel intense PA (mean over a long period of time) are the same people that feel intense NA at other times. The caveat of study 3 was the small sample of 26 students and that only a single time period was covered.

Study 4 had participants complete mood reports for six consecutive weeks as in study 3, summarizing the emotions of the day in the evening. Additionally, the participants had to complete mood reports at two random times every day. In this case they had to report the emotions they felt at this very moment.<sup>55</sup> While Diener and Emmons used no specific name for this method,<sup>56</sup> other publications used the term Experience Sampling Method (ESM) for this kind of methodology (Reed Larson and Csikszentmihalyi 1983). From now on, I will use the term ESM when describing questionnaires that participants

---

<sup>55</sup>The participants wore a watch with an alarm that they had to set according to a randomised list of times, hours in advance of the moment the rating had to be done. Participants were asked if they were thinking about the alarm. Most participants reported that in the first few days they were conscious about the alarm, but for the rest of the study period, the alarm had caught them unexpectedly. In later iterations of the experimental setup, it was taken care that every 5-minute period in the waking hours of a day had been covered by a questionnaire at least once (Ed Diener, Larsen, et al. 1985, p. 1258).

<sup>56</sup>The term "mood reports" was used for the questionnaires with the kind of items like "happy", "joy", "angry" etc. However, in this publication, the term "mood reports" would either refer to measures applied at the end of the day summarizing a whole day or mood detected at random times during the day. After this publication the setup where participants reported non-summarizing moods but the moods they felt that very instant at randomly chosen times during the day was used almost exclusively. Reed Larson who to my knowledge was one of the first to use the experience sampling method (ESM) with Mihalyi Csikszentmihalyi later collaborated with Diener on several publications.

had to fill out at random times during the day and respond about their mental status at this very moment, independently of whether Diener has used this term to refer to this method or not. Additionally to the daily summarizing and the moment reports, participants had to return three-weeks summarizing reports. Eventually they also reported on a separate questionnaire when they had felt very strong emotions, at maximum once per day. The correlations of the different time periods differed greatly. The strongest negative correlation in intensity between NA and PA within-subjects was at the emotional times. This correlation decreased with an increase of the time period covered. The shorter the time period, the stronger the negative correlation in intensity. With this study, Diener has shown that the independence of NA and PA in frequency and intensity is based on the length of the time-period covered. The change in methodology towards the ESM also ensured that memory-bias did not affect the results, since participants filled out the questionnaires immediately. Since all the previous studies participants had been undergraduate students, there was one last study to conduct before the paper could be wrapped up.

Study 5 participants were aged 33 to 85 and handed in day-summarizing mood-reports for 30 consecutive days.<sup>57</sup> The mood items were identical to study 3 and 4, albeit in enlarged print. The results confirmed the ones obtained from the college students. Separating the methodological steps in their own studies allowed to track every change in the method and its influence on the results. Getting results from two timeframes in the third group helped to establish that positive and negative emotions correlated more strongly negatively over the relatively short time period of a month than over a year. “The shorter the time period is, the more negatively correlated are the two types of affect” (Ed Diener and Emmons 1985, p. 1112). First Diener and Emmons changed the emotional items to ones they trusted and ‘validated’ them with Bradburn’s scale. After the graphemes had been calibrated the experimental setup of the ESM allowed to rule out memory bias completely. Through adaptation of their instruments, Diener and Emmons managed to get the data that confirmed their hypothesis and showed Bradburn’s hypothesis lacking.

The results of these five studies strongly supported the argument that the independence of frequency and intensity of NA and PA depends on the time period considered and offered a compelling explanation why previous studies using Bradburn’s instruments came to another conclusion (see *ibid.*, p. 1114). The research on affect frequency showed that a person’s positive mental state in a certain moment has the effect that this person’s negative mental state must be very low. “When a person is experiencing either positive or negative affect, he or she is less likely to experience the other type of affect” (*ibid.*, p. 1111). The frequency of NA and PA therefore is clearly negatively correlated short term. However, the intensity of either affect within-person seemed to correlate positively. Therefore, a person feeling one type of affect more intense on average will also feel the other one more

---

<sup>57</sup>Since study 5 only needed to rule out that the shown effects were student specific, the application of the expensive ESM was waived.



intensely – at another point in time. “Thus, mean levels of the two types of affect do not correlate when longer time spans are considered because the influence of frequency and intensity cancel each other out” (Ed Diener and Emmons 1985, p. 1116). The purported independence of NA and PA as detected by Bradburn was a measurement artefact by not differentiating between frequency and intensity and only taking long-term measures into account.

Each of the above-described studies built on previous work done, testing a new hypothesis and rejecting new antitheses and biases. While not described above, several studies in that paper also included response-style checks<sup>58</sup> and a questionnaire to check for answers based on social desirability. Further, study 3 and 4 also included some personality questionnaires. Some of the traits correlated with one affect positively and the other negatively, while other personality traits only correlate with one type of affect, most notably “extraversion” and “neuroticism”. These leads were not covered in study 5 but soon picked up in further studies.

The studies mentioned in this chapter were usually done building on each other’s results. The confirmation or adaptation of each working hypothesis then lead to a new set of hypotheses that were tested and if possible lead to a new research question. In the above-described study, an already existing instrument (the Bradburn scale) is compared to a new instrument that he had been assembled using previous research. After determining that the new instrument was working as intended compared with the old instrument, the researchers introduced further changes to the studies’ design to fix the perceived shortcomings of previous studies. In this case the possibility of memory bias and the impact of the inclusion of intensity of affect were checked. The first few studies were done on students of various psychology classes, leading to a limited age-range of participants and an over-representation of women in most cases. The gender of the participants was usually listed in the methods section. However, no other information, for example on ethnic background, was given about the participants,<sup>59</sup> other than the numerous reassurance that the “subjects were quite heterogeneous insofar as is possible in a student sample” (Emmons and Ed Diener 1985, p. 91). The methodology and hypothesis were tested and improved iteratively on the student population. Frequently, also in the above-described case, the last study in a publication confirmed the observed effects in a more general population that included people aged 33-85. The number of participants in this last study was usually limited, due to cost and since the goal of this previous study was not to add more data to the hypothesis, but simply to confirm the previous studies results. The instruments used in the last study were usually limited to the best working methods.

---

<sup>58</sup>These items checked whether certain people were prone to use big numbers for example when noting down their answers on the scales for example.

<sup>59</sup>With the exception when a cross-cultural sample was included. In the described paper the study 2 participants were students from the College of the Virgin Islands.

### 3.3 Frequency and Intensity of Affect

The purported influence of frequency and intensity on PA and NA was focused in the paper “Intensity and Frequency: Dimensions Underlying Positive and Negative Affect” published in 1985 (Ed Diener, Larsen, et al. 1985). The three studies in this publication followed the same general structure as in “The independence of Positive and Negative Affect”, where hypothesis were tested incrementally and the last study confirmed the results in a non-student population. There are two main points in this paper I’d like to focus on: artifacts checks and a change in methodology which lead to the definition of “frequency of affect”. So far participants had been filling out the questionnaires about their experienced affect. The results might therefore hypothetically be influenced by the way people see themselves and by their use of numbers. In order to check for the latter, participants had to rate 57 positive affect words on the level of intensity implicit in their meaning, not in how they felt at that moment.

“In this rating task any consistent tendency to use either high or low numbers had to be due to how people used the numbers of the scale and not the differences in their affective state because the same words served as stimuli for everyone” (*ibid.*, p. 1256).

If participants had tended to use high number when rating the intensity of these words, they could be expected to rate their felt affect stronger as well. When this intensity rating was partialled out in the results,<sup>60</sup> the correlation of intensity of NA and PA actually became stronger. Therefore, rating bias does not explain the use of numbers participants’ apply when filling out their mood forms.

In the second artifact check, self-report bias, should be rejected. Participants had to hand in three descriptions of times when they felt very happy and three descriptions when they had felt very unhappy. Each description had to be written within 30 minutes and participants were asked to write in great detail. Judges (researchers) then rated these text descriptions on affect, without knowing the participants’ score when measured by ESM. An estimate of affect was computed for each participant. The ratings of the judges, based on qualitative data and not influenced by the participants’ use of numbers, correlated well with the ESM self-reports of each participant. “[T]hese correlations represent strong evidence that the variance between subjects on the intensity dimensions reflect meaningful differences in how they experience affect” (*ibid.*, p. 1257). After years of asking people about their mental state, this study checked whether ‘questionnaire response behavior’ correlated with how people actually feel (as per their written description).

The last part that I want to discuss in this publication concerns a methodological trick Diener and his team applied. First, they chose 9 items that had to be reported on a

---

<sup>60</sup>This was done by partialling out the mean rating for the 57 positive affect words of the correlation between PA and NA (see Ed Diener, Larsen, et al. 1985, p. 1257).

monopolar scale based on factor analytic work. While items in previous studies were selected on the basis that there was a connection between them and NA/PA, these 9 were considered to best describe PA and NA.<sup>61</sup> The scores of these ESM reports were calculated in such a way, that they could all be put on a bipolar scale.<sup>62</sup> Putting NA und PA on the same scale made them directly comparable.

Bradburn had compared NA and PA purely on the level of frequency as expressed by the participants by memory. This new instrument had access to the frequency and intensity of NA and PA. Further, it did so based on ESM data that could not have been tampered by memory bias. Therefore, this new instrument allowed a summarized mood intensity and frequency report of the day that was based on actual data.<sup>63</sup> “The frequency of positive affect” would in this publication now be defined by “adding up the instances when the daily positive affect score exceeded the daily negative affect score, and dividing this number by the total number of days sampled. Therefore, the »frequency« dimension refers here to the frequency of predominantly happy days” (Ed Diener, Larsen, et al. 1985, p. 1256). As such, Diener and his team were able to produce a unified NA/PA score that was true to the discoveries that they had made on the journey.

Diener and his team concluded that the intensity of affect is a vital personality variable. However, it is not a personality variable only. Just about half of the variance in intensity score seemed to be accounted for by this underlying personality factor. The rest of the intensity of PA and NA must be influenced by unknown factors, probably the context of the situations. The strong positive correlation of intensity of PA and NA and the strong negative correlation between the frequency of PA and NA seem to cancel each other out over time, hence the claimed independence of PA and NA by Bradburn and others. With these studies, Diener and his team were able to explain why some scales show that PA and NA are independent and others suggest that they are inversely related, depending on whether their wording taps mean levels of intensity or frequency. The introduction of better and more sophisticated instruments lead to unmasking one of the most important empirical findings concerning the relationship of NA and PA to be a statistical artifact, while at the same time explaining why previous instruments confirmed the purported independence of NA and PA.

### 3.4 The Personality Structure of Affect

I would like to present two additions to the toolbox in the paper ”The Personality Structure of Affect” (Ed Diener, H. Smith, and Fujita 1995): To the already introduced day-

---

<sup>61</sup>“Happy,” “pleased,” “joyful,” and “enjoyment/fun” represented PA and “unhappy,” “depressed/blue,” “frustrated,” “angry/hostile,” and “worried” represented NA (Ed Diener, Larsen, et al. 1985, p. 1256).

<sup>62</sup>Four terms were computed into a daily positive affect score and five for a daily negative affect score. The PA and NA scores had to be divided by the number of terms per affect in order to be comparable. Therefore, the PA score was divided by four and the NA score was divided by five. If the total PA score was higher than the total NA score, said day was considered a PA day.

<sup>63</sup>The data was not summarized by memory by the study participants.

summarizing self-reports and scales that only had to be filled out once (i.e. personality scales).

The first new instrument I would like to introduce are the so-called peer-reports. In the peer-reports, sometimes also called informant-reports, friends and family filled out questionnaires for how frequently they perceived specific emotions of the participant to be over the past month.

*Factor Analyses of Adjectives Within the Discrete Emotion Categories*

Basic category and emotion words	Factor loading	Percentage of variance
Love		70.0
Affection	.86	
Love	.87	
Caring	.86	
Fondness	.75	
Joy		62.1
Joy	.84	
Happiness	.87	
Contentment	.75	
Pride	.68	
Fear		68.2
Fear	.72	
Worry	.84	
Anxiety	.86	
Nervous	.87	
Anger		66.6
Anger	.83	
Irritation	.82	
Disgust	.79	
Rage	.82	
Shame		64.3
Shame	.82	
Guilt	.85	
Regret	.74	
Embarrassment	.80	
Sadness		64.7
Sadness	.77	
Loneliness	.76	
Unhappiness	.84	
Depression	.85	

Figure 3.2: Factor Analyses of Adjectives Within the Discrete Emotion Categories (Ed Diener, H. Smith, and Fujita 1995, p. 134).

By using four completely different instruments (daily reports, summarizing the day, peer-reports summarizing the month and one time self-report summarizing the month) the researchers hoped to evade the most constant sources of errors since it is very unlikely that an error enters all four measures at once. The emotional terms in the peer-reports could be grouped into a total of six discrete emotions.

Gathering data by so many different instruments allowed using the second tool portrayed here, the Multitrait-Multimethod Matrix (MTMM), developed by D. T. Campbell and Fiske (1959). In MTMMs measures of the same trait by different methods should be seen to have higher correlation than measures of different traits measured by different methods. Usually, correlations of same trait, different methods measures

are also expected to be higher than measures of different traits by the same method. As seen above, every discrete emotion was represented by four adjectives in the questionnaires. These four adjectives were now summed within each method and compared to one-another. The three methods produced very similar relative frequencies of the emotions encouraging the researchers to use the data to establish how often which emotions are felt. MTMMs are useful to compare the same emotions in different methods (monotrait-heteromethod) showing the degree to which the same constructs converge using different methods (convergent validity).

**Multitrait–Multimethod Matrix for the Positive Emotions**

Type of report and emotion	Self-report		Daily		Informant	
	Love	Joy	Love	Joy	Love	
Self						
Love	$\triangle$					
Joy	<u>.63</u>					
Daily						
Love	<u>.65</u>	.45	$\triangle$			
Joy	<u>.48</u>	<u>.68</u>	<u>.70</u>			
Informant						
Love	<u>.54</u>	.34	<u>.55</u>	.34	$\triangle$	
Joy	<u>.41</u>	<u>.45</u>	<u>.40</u>	<u>.52</u>	<u>.62</u>	

*Note.* Underlined values represent convergent validity

Figure 3.3: Multitrait-Multimethod Matrix of PA (MTMM) (Ed Diener, H. Smith, and Fujita 1995, p. 134).

Since heterotrait-heteromethod correlations are smaller than heterotrait-monomethod correlations, a slight method bias can be observed. Heterotrait-heteromethod correlations in adjacent rows and columns also showed greater correlations between discrete emotions than across summarized NA and PA. Since a single method produces greater convergence than across methods, the researchers stressed

the influence of method variance on emotion measures and the importance of using different methods. With the correlation matrix of the six discrete emotions times three different methods the researchers proposed a model of affect structure.

The model shows that while the correlations between the respective emotions and PA/NA were confirmed as expected, the emotions could not be conflated into global PA and NA groupings. Several emotions correlated with a specific other emotion in a way that was not accounted for by global PA and NA.

### 3.5 Re-Assessing Frequency of Affect

As I have tried to show, new studies aimed at providing data for a particular hypothesis usually incorporated old methodology which's behaviour was known, as well as new methodology which should produce better data to support or discard the new hypothesis. In order to advance knowledge on SWB, Diener required new methods which first had to be validated through convergent results of the old methods. The results of a new measure have to correlate with old measures to make sure it picks up on the same phenomenon and delivers new data to be used for the new hypothesis.

In 1991 Diener and his colleagues published a book chapter with the telling title “Happiness is the Frequency, Not the Intensity, of Positive Versus Negative Affect” (Ed Diener, Sandvik, and W. Pavot 1991). In it, the authors discuss the importance of frequency and intensity

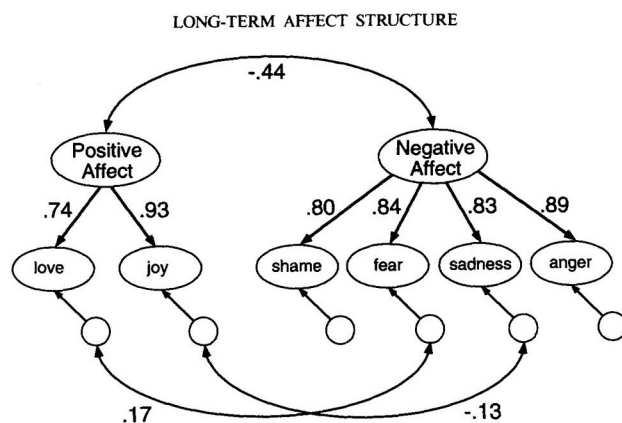


Figure 3.4: Two-factor model of affect structure by (ibid., p. 137).

of affect for SWB and make the practical argument that one of the main reasons to focus on frequency over intensity is that it “is more easily and accurately measured than affect intensity” (Ed Diener, Sandvik, and W. Pavot 1991, p. 120) Number of PA and NA instances are easier to measure and differentiate than several degrees of intensity. Diener, Sandvik and Pavot held that the two basic emotional experiences of PA and NA are largely universal, whereas intensity categories on scales from one to seven are probably subjective, given intensity values correlate positively within-person. Previous studies had also shown that participants were pretty accurate in remembering frequency<sup>64</sup> from memory, while affect intensity reports from memory were heavily skewed (Thomas and Ed Diener 1990). Further, happiness/unhappiness is easier to code within and across people. Most people can tell whether they were happy or unhappy at a particular point in time, whereas there is “no natural system by which to define or label emotional intensity” (Ed Diener, Sandvik, and W. Pavot 1991, p. 121).

The studies in “Happiness is the Frequency, Not the Intensity, of Positive Versus Negative Affect” used the methods described-above adapted to the questions posed. Diener and his team were able to establish that *intense* PA is neither necessary, nor sufficient for being happy. The *frequency* of PA was always the better indication of high SWB rating than the *intensity* of PA was.

“Subjective well-being measures all converge on the property of frequent positive affect, indicating that this experience is the essence of a phenomenon which can be labelled »happiness«. [...] Thus, we maintain that happiness should be defined as relative frequent positive affect and infrequent negative affect” (*ibid.*, pp. 123–124).

And thus, the research group around Diener has arrived at an operationalization of ‘happiness’: Living a life, in which over a certain period of time, more days pass where more emotionally positive instances happen than negatively read ones. From an instrument and operationalization point of view it is interesting that this operationalization does not rely on one measurement over an extended period (memory bias!) but on several measurements over an extended period of time that then can be summarized into *one* score.

### 3.6 The Iterative Process of Studies and Methods

Diener and his colleagues made many adjustments to the methodology not in order to get data for the hypothesis but for being able to exclude alternative hypotheses and minimise possible bias. The new studies and instruments made sure the data is reliable. The researchers feared that memory bias could influence the self-report measures. For those

---

<sup>64</sup>As a reminder, “frequency” of PA as it was measured is reporting more days where the total number of PA experiences excel the days where more NA events were experienced. Given a more accurate measure, frequency of Positive Affect is “the overall percentage of time the person is in a predominantly positive (as opposed to negative) emotional state” (Ed Diener, Sandvik, and W. Pavot 1991, p. 120).

reasons, the participants of the long-term studies had to turn in their forms daily in order to enforce that the participants had filled out the questionnaires as they were supposed. Participants failing to do so were excluded from the study. A further possible artefact of self-reports is differences in participants' number use. The scores of the so-called "response styles" were partialled out of several measures of affect in the studies but did not change the correlation between PA and NA intensity. The correlation being an artefact of self-report measures became more unlikely through the result of these precautionary tests (Ed Diener, Larsen, et al. 1985, p. 1259).

A further technique employed to detect potential artefacts was to use different methodologies for the allegedly same construct to measure. The most used methods were daily-summarizing self-reports, in-moment ESM-reports and peer-reports. Less used methods were texts that were rated by the researchers and structured interviews. All these methods could be adapted in order to best answer the questions of the studies by changing the time-frame about which participants had to respond or by changing the "trigger" for filling out the questionnaire (evening, random and at highly emotional times). As the data showed, the correlations between positive and negative affect intensity for the two emotionally most intense were even stronger than those for the summarised daily reports at the end of the day (*ibid.*, p. 1259). These results support the thesis that intensity is partly a character trait and has less to do with the situation the participant is in. Frequency correlation was strongly inverse when intensity was partialled out.

The studies were ordered in an iterative process. The first study had students fill out mood forms in the evening, establishing the effect of intensity and frequency. The second study added random time mood reports by ESM and the high emotional time mood reports while keeping the daily reports in the evening. Since Diener and his colleagues had tested the methodology in the first two studies on students, the third study had the sole purpose of replicating the data in an older population. The researchers deemed the data similar enough to be able to generalise the results of the first two studies.

These studies have shown that the behaviour of the three hallmarks of SWB was to this point in time still mostly unknown. Furthermore, the studies have uncovered that mood intensity might be partially caused by a character trait and not by the situation a person is in. It seems fitting then, that Diener and his PhD students became interested in studying the influence of character traits next. The correlation between affect and certain character traits had already been established (P. T. J. Costa and McCrae 1980)<sup>65</sup>. Applying factor analysis to the results of a study using several questionnaires showed two main factors of character traits when paired with Bradburn's scales: extraversion and neuroticism.

While extraversion correlated with PA, neuroticism correlated with NA. Several other character traits had been found that correlate with NA and PA in roughly inverse ways,

---

<sup>65</sup>Paul Costa and Robert McCrae were two of the very few researchers who used the term "subjective well-being" before 1984 in the sense as Ed Diener does from 1984 on.

like self-esteem and self-confidence. Extraversion and neuroticism were chosen because they tend to only correlate with one affect and not with the other (Ed Diener, Larsen, et al. 1985, p. 1259). In the paper "Personality Correlates of Subjective Well-Being", Emmons and Ed Diener (1985) compared long-term mood reports in the evening of each day with SWLS results and several questionnaires measuring personality dimensions. Scales measuring extraversion correlated highly with PA, while neuroticism ratings correlated with NA. Variables correlating with PA did not correlate significantly with NA and vice versa. Emmons and Diener concluded that "individuals who experience much positive affect typically are extraverted (sociable but not impulsive), active, enthusiastic [...]. Individuals for whom negative affect predominates are emotionally reactive, interpersonally sensitive, [...] are low in self-esteem and tend to believe that outcomes are not under their own influence" (Ed Diener, Larsen, et al. 1985, p. 94). As such, the paper supports the thesis of P. T. J. Costa and McCrae (1980) added more detail to it, while also showing that extraversion-related personality traits correlate with life satisfaction, as measured by the SWLS (Ed Diener, Larsen, et al. 1985, p. 92).

A follow-up study where participants described the most intense events per day in a text and rated it on a scale showed a clear difference between high-intense subjects who responded more strongly to life events and less intense subjects who responded more moderately. "Furthermore, this effect held no matter how objectively bad or good those events were" (Larsen, Ed Diener, and Emmons 1986, p. 808). A second study tested the reaction of students to hypothetical events. A group of students listed events they deemed typical for their everyday life. This ensured that the questionnaire items were close to what the students experienced in their own life. The items on the questionnaire had also been rated by a large number of students on how good or bad the events were. Another group of students now rated how they would react to said events. The comparison of the results of how good/bad the events were to how participants of the study would react to the imagined events mirrored the results from the first study. High-intense subjects gave more intense subjective emotional responses in the positive and in the negative spectrum on the fictional events, regardless of the severity of the described situations. The findings of this study rule out the possibility that the prior measured differences of high-intensity subjects and low-intensity subjects are due to a different lifestyle. As was established before with the text report methodology, self-report measures do work in the sense that participants do report the intensity of events in the same direction in text as in number form on a scale. As the last publication showed, however, the intensity reported for the same events may differ from person to person. Certain people tend to perceive events more intensely than others independently of whether they had experienced the situation or not.

Previous studies had established a link between positive affect and extraversion. Ideas, why this might be the case, ranging from the influence of more social contact of extraverts to a better person-environment fit since most people are embedded in societies with a lot



of social interactions. Some models of personality suggested that people that are high in anxiety have a greater sensitivity to signals of non-reward. Conversely, it could be possible that extraverts are more sensitive to signals of reward, lessening the impact of non-reward or punishment signals. While Diener and his colleagues expected extraverts to find such rewards in social interactions mostly, they conceived it believable that extraverts would find other rewards in the environment in the absence of social interactions. “This approach suggests a temperamental predisposition for extraverts to experience more happiness, regardless of the social features of the situation” (B. Pavot, Ed Diener, and Fujita 1990, p. 1300).

The memory difference measure was another non-traditional measurement that was used in this study. This method was devised based on observations of other research groups. They had observed that people tend to report memories that are congruent with the mood at the time of recall. In the memory difference measure participants had to recall as many possible memories within 3 minutes and as many possible unhappy memories within another 3 minutes. Since happy people were expected to have stronger cognitive network for happy memories this test should check whether they have a positive memory bias. The difference in the number of events within the two intervals has been used as an additional measure of global SWB and was validated through convergence with other measures of SWB (Ed Diener, Sandvik, W. Pavot, and Gallagher 1991).

Additional to the self-reported data, family and friends of the participants had to give peer reports on the SWLS for the study participants in order to calculate the participants' SWB. In all these measures, the gathered data offered little evidence to suggest that the greater amount of time spent in social contact with others leads to the higher levels of SWB as measured in extraverts. The ESM reports did not show that extraverts were significantly more in social situations than introverts were, they did report higher mean levels of PA though. The results also suggested, that more PA is experienced in social situations and greater levels of NA when being alone. The described study did show the cross-method robustness of the extraversion/neuroticism to SWB relationship. Further, this robustness was demonstrated with three non-self-report methods (rated text, peer reports, memory measure), all of which correlated considerably with self-report measures.

Diener and colleagues conducted a four year-long study in order to test whether personality traits were stable and whether extraversion would lead to more reported positive life events. At the same time, neuroticism was expected to lead to more reported negative life events (Magnus et al. 1993). Taking the same personality, mood and happiness tests twice with a time gap of four years showed that the neuroticism and extraversion scores were stable for the majority of people. Based on existing scales, a life events list was constructed and adjusted to the student population aged 22-26. Another group of students had to rate whether the items were objectively either good or bad so that only clearly categorisable events like “death of a close family member” or “got into graduate school” were on the

list. A second group of students then rated the life events on positivity and negativity. The list of the most positive and most negative life events were then presented to the participants of the four-year study, where they indicated which life events had actually happened to them within these four years. Participants high in extraversion experienced a greater number of objectively positive events, whereas participants high in neuroticism experienced more objectively negative events. While the correlation of extraversion with positive affect and neuroticism with negative *affect* had prior been established (P. T. J. Costa and McCrae 1980), this study showed the expected relationship of extraversion and neuroticism with *life events*. “Extraversion predisposed participants to experience a greater number of objective positive events, whereas neuroticism predisposed the occurrence of objective negative events” (Magnus et al. 1993, p. 1050). Prior studies let participants decide themselves if the experienced life events were positive or negative. Comparing this study with previous studies unveiled that reports of subjectively rated life events do not show a greater relation to personality than objectively rated life events. The study also made clear that the influence of life events on life satisfaction depends on personality to a certain extent. Therefore, when checking the influence of life events, personality traits also have to be considered (*ibid.*, p. 1052).

Why participants higher in SWB experience more positive events was to be partially uncovered by the following study: Participants completed several measures, including the SWLS, current mood and the memory difference measure, wherein they had to recall positive and negative life events within a set amount of time (Seidlitz and Ed Diener 1993). The study showed correlations between remembering positive and negative life events, mood and SWLS. A second study was to clarify two hypotheses: 1. The recall difference stems from the difference between happy and unhappy people. 2. Recall differences are based on subjective interpretations of life events. Participants responded to SWB measures and did the memory recall test. Based on SWB measures a “negative group” and a “positive group” was formed. The participants received a pre-assembled list of life events. They had to mark those events that they had recalled during their memory recall test and then rate these events by their favourableness. The events on the list had already been rated objectively by other students. Comparing the subjective ratings of events that had actually happened to the participants with the objective ratings showed that “differences between happy and unhappy subjects in recall of positive versus negative life events resulted from differences in the incidence of life events and from differences in the interpretation of events” (*ibid.*, p. 662). Furthermore, the “positive group” recognised more positive versus negative life events on the checklist as having actually happened to them. This result again stressed that simply remembering positive and negative life events is not an accurate measure since personality traits influence the interpretation of the affect tied to the events.

### 3.7 Cultural Influence

Most of the studies Diener and his colleagues had done to this point had been administered to Americans and American students specifically. Several studies were conducted to check if the measures, when applied to a different population, would produce similar results. For the researchers, the most crucial point of such studies would not be which score a different population reports, but if SWB is structured in the same way as has been concluded on the American population. The factor analysis of the results of a study on Russian students showed that results from the SWLS formed a unitary factor, similar to the results in the U.S.A. This was an encouraging result as the SWLS, therefore, seemed to measure life satisfaction and life satisfaction only, also when applied to a different population. Also, the administration of the D-T Scale about life-as-a-whole formed the same three factors (life satisfaction, PA and NA) as Andrews and Withey (1976) had found in the U.S. population. Even though the scores of the SWLS and the D-T Scale about life as a whole showed that Russian students were in generally less satisfied with their lives than their American counterparts, the structure of these and further measures in November and December 1990 led Diener and colleagues to conclude that the facets of SWB are consistent across cultures (Balatsky and Ed Diener 1993). The researchers also noted, however, that there were marked differences between “individual” and “collectivist” cultures which influence the importance of certain life domains on the SWB rating (Ed Diener, E. M. Suh, H. Smith, et al. 1995). Other studies by Martin Seligman and Yueh-Ting Lee showed that in a Chinese population failure and success was attributed more often to other people, which resulted in less self-blame (Lee and Martin E.P. Seligman 1997). While this result supported Seligman’s “explanatory style” theory on the influence of internal and external explanations, the results also support Diener’s hypothesis on the difference of certain life domains on SWB in individual and collectivist societies.

### 3.8 Putting Things Together

Reviewing his research in 1994, Diener wrote that definitions of SWB were often not made explicit but only implied by the measures used. A composite definition of SWB suggests three hallmarks to the area of SWB and has been left mostly unchanged since 1984: First, SWB is subjective; the experience resides within the individual. Second, SWB is not solely the absence of negative factors but has to include positive measures. Third, a measure of SWB has to include global (whole life) assessments and must not focus on single life domains (Ed Diener 1994, p. 106).

“The implicit theory of SWB [...] is the following. Humans are not only capable of appraising events, life circumstances, and themselves, but they make such appraisals continually. Appraisals of things in terms of goodness-badness is a human universal. [...] Such appraisals are seen as leading to emotional reactions,

which can be either pleasant or unpleasant. Other things being equal, pleasant experiences are perceived as desirable and valuable. Thus, a person who has pleasant emotional experiences is more likely to perceive his or her life as being desirable and positive” (Ed Diener 1994, pp. 106–107).

While it has been implied through the measurements but seldomly explicitly mentioned, SWB researchers “are primarily interested in factors which lead to specific levels of well-being over periods of weeks, months, or years” (*ibid.*, p. 109). So, to count as a ‘good’ measure of SWB, not only have several instruments to converge on the result, the result also has to pick up on a ‘stable’, long-term phenomenon. On the self-report and non-self-report methods, Diener and colleagues argue that the measures show great convergence and contain substantial amounts of common variance that points to a coherent phenomenon. Diener refers to a study by P. T. Costa, McCrae, and Zonderman (1987) showing demographic factors like health, income and physical attractiveness having surprisingly little correlations with SWB. In comparison to demographic variables, personality variables like extrovertedness, introvertedness and other factors like unemployment seem to have a stronger effect. Diener refers to the unemployed as the unhappiest group the happiness-scholars have studied so far.

As explained in the studies further above, the observed independence of NA and PA by Bradburn and others depended mainly on the time frame, the intensity of the emotions sampled and whether verbal or nonverbal measures were used (Ed Diener and Larsen 1993, p. 407). By using the ESM and not day-summarising measures, Diener and his team could show, that participants did not report NA and PA simultaneously. Had NA and PA been truly independent, participants should have reported this phenomenon. Short-term measures showed a clear inverse (negatively correlated) relationship between NA and PA. When long-term measures were applied that also included intensity measures, this relationship began to blur. While Diener argues for an inverse relationship of NA and PA under the conditions mentioned above, he still holds that they have to be measured independently since factor analysis shows them to be two discriminant factors.

### **3.9 The State of SWB Research at the Advent of Positive Psychology**

Demographic factors are not part of the standard SWB measures, since studies have shown that they are weakly correlated with SWB (see Andrews and Withey 1976; A. Campbell, Converse, and Rodgers 1976). At the same time, some life situations are known to be influential. Studies have shown differences in SWB between children in intact and divorced marriages (Gohm et al. 1998), while conclusions on the impact of wealth are conflicting. While SWB reports in the U.S.A., Japan and France have not changed over 20 years in which more income has been gained (Ed Diener and Ed Diener 1995), people whose income increased or decreased did not change SWB (Ed Diener, Sandvik, Seidlitz, et al. 1993), while differences in SWB reports in poor and rich nations are substantial. In general,

Diener holds that “external circumstances are often less important to SWB than is often believed, probably because people partially adapt to them” (Ed Diener, E. Suh, and Oishi 1997, p. 38). Right around the time of the foundation of positive psychology Diener and his colleagues published an overview paper called “Subjective Well-Being: Three Decades of Progress” (Ed Diener, E. M. Suh, Lucas, et al. 1999). Even though SWB research had picked up the pace, psychological articles examining negative states outnumber articles exploring positive states by a ratio of 17 to 1 (Myers and E. Diener 1995). On the state of research, Diener notes:

“SWB research is limited by the almost exclusive reliance on cross-sectional correlational designs with inadequate tests of causal hypotheses. This shortcoming leaves researchers in an uncertain position regarding the causal priority of the variables they study” (Ed Diener, E. M. Suh, Lucas, et al. 1999, p. 277).

Here, causal modelling and longitudinal studies have been conducted to help but need to be improved. Since objective circumstances seem to influence SWB only on a small level, some researchers believe that ‘happiness’ is a character trait. Since stable personality traits like extraversion and neuroticism can influence SWB however, Diener and colleagues hold that SWB has both trait-like and state-like properties.

“The current working model of researchers in the field is that personality predisposes people to certain affective reactions but that current events also influence one’s current levels of SWB. [...] However, long-term life circumstances can also have some continuing influence on people’s level of SWB” (*ibid.*, p. 280).

The reason for the correlation with extravertedness and neuroticism is still unclear, but “an intriguing possibility is the idea that the characteristics of extraverts are actually an outcome of higher levels of positive affect” (*ibid.*, p. 281). While initially the research on SWB based a lot of the arguments on observations by members of the social indicators movement, the approach in the field has changed: the graphemes chosen for the reasons outlined above, produced data that within this vast experimental system suggested that a such operationalized ‘happiness’ as subjective well-being, is a rather internal matter.

## Chapter 4

# The Rise of Positive Psychology

### 4.1 Short Introduction to Seligman's Work and Theories Before 1998

The aim this chapter is to describe the different approach that Martin Seligman applies to the research on happiness when compared with the field of Subjective Well-Being. In the last chapter I argued that the field of SWB as observed in the research group around Ed Diener seemed to be mostly pushed forward and constrained by what is observable by measurements. Ed Diener defended the lack of theory building in the field with the argument that he wanted to secure some facts first before a theory can be formulated. Martin Seligman started the new movement of "Positive Psychology" in 1998 in his American Psychologist Association presidential address and focuses on the possibility of interventions (see Martin E. P. Seligman 1999b). In this chapter I argue that this change of focus has serious implications for the theory of 'happiness' formulated within these differently formulated experimental setups.

Martin Seligman's academic career started with a PhD in the experimental psychologist lab of Richard Lester Solomon at the University of Pennsylvania in 1967. In his experiments with dogs and rats Seligman noticed a curious behavior which he later coined "learned helplessness": Two groups Dogs were put in a "shuttle box" which consisted of two compartments with an obstacle in the middle. In the first stage, the first group of dogs received inescapable electric shocks administered through the metal grating floor. The obstacle in the middle of the shuttle box was raised during this stage of the experiment. The second group of dogs did not receive this treatment. In the second stage of the experiment both groups of dogs were warned of the impending electric shock with a sudden dimming of the light in the shuttle box. The barrier was lowered, so the dogs could escape the electric shock by jumping over it. Once they had learned to heed the light cue in the second stage, some of dogs were able to completely avoid the shock by moving to the safe side of the box. Seligman expected the first group of dogs quickly to pick up on the cue and escape to the other side of the shuttle box since these dogs had experienced the shock before. The second group who had not experience the shock prior to the experiment was expected to be slower to react. The opposite was the case. The first group of dogs mostly

disregarded the cues and endured the shocks while the second group managed to escape.<sup>66</sup>



Figure 4.1: The "shuttle box" within which the dog experiments took place (see Overmier and Leaf 1965; Martin E. P. Seligman and Maier 1967), illustration from (Spielman et al. 2017, p. 519).

After rejecting several unsatisfactory hypotheses for the effect, Seligman and his co-author Steven Maier concluded that dogs in the first group had learned to be helpless and had internalized by the first stage when they were given inescapable electric shocks, that they could not help themselves. In the second stage of the experiment they were unable to read the new situation properly. This experiment had convinced Seligman of the flaws of behaviorism, since only taking the point of view of cognitivism and assuming that the mental states of the dogs had caused their behavior allowed him to explain the experimental results (see Martin E. P. Seligman and Maier 1967). Soon after, Seligman interpreted the experiments results towards an explanation of humans with depression and was looking for ways to improve psychotherapy (see, for example Abramson, Martin E. Seligman, and Teasdale 1978; Martin E. P. Seligman 1972; Maier and Martin E. P Seligman 1976). The 1982 published "Attributional Style Questionnaire" (ASQ) (Peterson, Baeyer, et al. 1982) was an instrument based on the "Learned Helplessness" theory which again was built on the basic conclusions of the dog experiments: it is not the events themselves,

<sup>66</sup>This is a summarizing illustration of the experiments with a focus on the main points that lead Seligman to his conclusions on learned helplessness. This actual study covered three experiments with up to four groups of dogs each, whereas the main conclusions were drawn in a second study that was able to reject more alternative explanations that would have been in line with behaviorism (see Overmier and Martin E. P. Seligman 1967; Martin E. P. Seligman and Maier 1967).

but the internal explanation of the subject that are the determinants whether someone was depressed or developed a depression. The Attributional Style Questionnaire works as follows: Participants are presented with one of several good or bad imaginative situations<sup>67</sup> and can choose among the following explanations why such a situation could happen to them: They might experience the event due to "other people", "circumstances" or "me". Further, participants are asked whether they expect this cause will "always" or "never again" be present when the same event like looking for a job will present itself in the future. The last question presented here has participants answer whether they expect the cause of the event influencing only this type of event or other areas of life as well. The Learned Helplessness theory predicts that an explanation style "in which internal, stable and global attributions are offered for bad events is associated with depressive symptoms" (Peterson, Baeyer, et al. 1982, p. 296), *i.e.* not getting a raise because of other people, expecting the same to happen again for the same cause and this cause to influence other areas of life. Learned Helplessness theory proposed the definition of depression describing it as a result of experience with uncontrollable aversive events. Like the dogs in the experiments, depressed participants concluded that they were not in control of the situation and could not help themselves. Peterson and Seligman found that when ASQ scores of students were taken after midterm examinations, ASQ scores were correlated with depressive symptoms following poor exam performance. Further studies confirmed ASQ scores predict which study participants were to develop depressive symptoms one month later (Golin, Sweeney, and Shaeffer 1981).

Later, Attributional Style theory underwent a name change and was called "Explanatory Style". However, only the theory underwent a name change, the attributions of the explanatory style theory were still measured with the Attributional Style Questionnaire instrument. With this name change came an extension of application of the new explanatory style into the 'positive' mental realm. Explanatory style, as measured by ASQ, was not only able to predict how people *felt* but also how they *performed*. Seligman argued with explanatory style that explaining bad situations by one's own agency is having negative effects on overcoming future challenges, while explaining good situations being brought by through one's own agency correlates with doing good in the future, whereas 'doing good' was defined pretty broadly. By applying applying the ASQ to a population of newly recruited insurance sales agents was able to predict their job performance. The top half of the agents as measured by ASQ scores sold 37% more insurances in their first 2 years of service than the agents scoring in the bottom half of ASQ scores (see Martin E. P. Seligman and Schulman 1986). On the other hand, students with a bad ASQ score were more prone to illness years after the test (see Peterson and Martin E. P Seligman 1987). Both of these studies were seen as confirmation of the ASQ and explanatory style theory.

Interesting on a methodological level is that when Peterson, Baeyer, et al. (1982) published

---

<sup>67</sup>Unsuccessfully looking for a job, getting a raise, getting rejected, getting very rich etc.



the Attributional Style Questionnaire, they had not included a stage-like build-up of pre-studies to confirm their working hypotheses within the paper as shown in Diener's studies. The authors do mention some pre-studies, like testing questionnaire items on students in footnotes (*i.e.* Peterson, Baeyer, et al. 1982, p. 290), but no data is shown. The authors did discuss validity with reference to other authors having correlated low ASQ to depression (see Golin, Sweeney, and Shaeffer 1981). They further mentioned having replicated this correlation several times with students, but did not add a reference. So it must be assumed that these replication results have never been published. Further, previous studies by the authors of "Learned Helplessness" were mentioned, giving good reason that the ASQ will work in the predicted way. My interpretation is that the ASQ was published to the scientific community based on good arguments and tested on 130 undergraduates. The main difference to Diener-like studies lies mostly in the transparency of the scientific process. Unlike in Diener's studies, pre-studies were not included in the paper, neither were their data or methodological setup disclosed.

Until around the end of the 1990s most of Martin Seligman's articles revolved around depression, bad health and how to prevent it. However, two of his books that were written for the general public already carried in them what would later become positive psychology. *Learned Optimism* (Martin E. P. Seligman 1991) and *The Optimistic Child* (Martin E. P. Seligman, Reivich, et al. 1995) argue for the value of optimism beyond resilience against depression but stress a self-therapy approach on how to lead a better and more successful life.

## 4.2 The Founding of Positive Psychology

In 1998 Seligman got elected as president of the American Psychological Association. In his presidential address he called for a new science of human strengths that aims to identify the characteristics of a positive life, develops tools to measure and teach them (Fowler, Martin E. P. Seligman, and Koocher 1999). In the following paragraphs I'd like to stress some of the characteristics with which Seligman describes this new science and hint at the basic assumptions and approaches with which this new science started. The "Positive Psychology Manifesto", was written with other authors that have worked in the field of happiness-research, among others with Barbara Fredrickson, the creator of the "broaden-and-build model", and Mihaly Csikszentmihalyi, the creator of the "flow" theory. The goal of positive psychology was defined to be able to provide response to adversity and the study of how to bring about "the good life"<sup>68</sup> (K. Sheldon et al. 2000, p. 1), without further defining what "the good life" is. We will come back to Barbara Fredrickson and Mihalyi Csikszentmihalyi later in this chapter. In the introductory chapter of the special issue of the *American Psychologist* on positive psychology, Martin Seligman and Mihalyi Csikszentmihalyi criticise general psychology for turning a blind eye on what

---

<sup>68</sup>In quotes in the original.

makes life worth living, humanistic psychology for not building a cumulative empirical base and philosophy for dubious assumptions on what happiness is, without actually measuring anything (see Martin E. P. Seligman and Csikszentmihalyi 2000). From all these introductory texts in the new positive psychology we can infer that the aim was to focus on empirical work, while having therapy and interventions in mind.

### 4.3 A Definition of Happiness: The Three Lives

While Seligman gives a rudimentary formula for happiness in his bestseller *Authentic Happiness* (Martin E. P. Seligman 2002), I prefer to focus on the description given in *Positive Psychotherapy* (Martin E. P. Seligman, Rashid, and Parks 2006), since it is a more detailed and scientific description and publication in general. Seligman suggests, that “the unwieldy notion of ,happiness‘” (*ibid.*, p. 776) can be decomposed in three parts: The pleasant life, the engaged life and the meaningful life.

#### 4.3.1 The Pleasant Life

Seligman’s understanding of the pleasant life is based on hedonic theories of ‘happiness’ whereas positive emotions about the present, past and the future work as a buffer against depression and anxiety (*ibid.*, p. 777). As a sidenote I must add that Diener’s SWB (based on a hedonic definition of happiness) and Seligman’s idea of hedonic definitions of happiness have some marked differences. Seligman describes the pleasant life as “having a lot of positive emotion [...] and learning the skills to amplify the intensity and duration of these emotions” (*ibid.*, p. 776). Diener would vigorously disagree about the intensity of the emotions and stress that it is the frequency of said emotions that matters, while referring to his multistaged studies, showing that high intensity of positive emotions often comes at the cost of having had low intensity of positive emotion for a long time beforehand (see Ed Diener, Sandvik, and W. Pavot 1991). As Diener argues, higher measures of life satisfaction as measured by the SWLS are a product of frequent, but not necessarily intense measures of PA. In consequence, an amplification of ‘happiness’ understood in the reading of the proxy construct subjective well-being is not the product of the amplification of the positive emotions themselves (aka their intensity), but their frequency.

In his definition of the pleasant life Seligman relies mostly on Barbara Fredrickson’s broaden-and-build model (see Fredrickson 1998; and Fredrickson 2001). The broaden-and-build model follows the observation that negative emotional states like anxiety and fear narrow people’s focus, while high-arousal positive emotions broaden attention and focus. Barbara Fredrickson and Thomas Joiner used the Positive and Negative Affect Schedule (PANAS) questionnaire (see D. Watson, Clark, and Tellegen 1988) and the “Cognitive Response Inventory” (CRI) to show that measures of positive affect (but not negative affect) are able to predict broad minded coping when confronted with a complex problem

5 weeks after taking the measure (see Fredrickson and Joiner 2002). The PANAS was used to measure the intensity of some emotional items by memory over the past two weeks, while in the CRI participants have to remember an important problem within the last year and indicate how often various coping strategies have been used. Fredrickson and Joiner further were able to confirm their hypotheses that PA and broad-minded coping serially enhance each other. In essence, the broaden-and-build model proposes an upward spiral of PA and broad-minded coping, where the two variables positively influence themselves and one-another.

Seligman further refers to Sonja Lyubomirsky, Kennon Sheldon and David Schkade's paper "Pursuing Happiness: The Architecture of Sustainable Change" (Lyubomirsky, K. M. Sheldon, and Schkade 2005) in order to argue for exercises that enhance positive memories and therefore the pleasant life. The authors definition of happiness reads: "frequent positive affect, high life satisfaction and infrequent negative affect" (*ibid.*, p. 115) and is explicitly based on Diener's work. By applying Diener's SWLS, the PANAS and other measures, Lyubomirsky and colleagues were able to confirm that an intentional change in one's activity has a longer lasting influence on one's happiness than circumstantial change, therefore arguing that "effort and hard work offer the most promising route to happiness" (Lyubomirsky, K. M. Sheldon, and Schkade 2005, p. 116; and K. M. Sheldon and Lyubomirsky 2006).

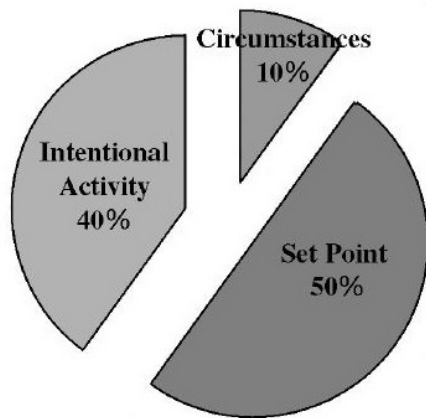


Figure 4.2: The three primary factors that determine "chronic" happiness levels (*ibid.*, p. 116).

The "chronic happiness" in figure 4.2 is defined as "a person's characteristic level of happiness during a particular period in [...] life (*ibid.*, p. 115). Lyubomirsky and Sheldon's studies in turn rely on the data of David Lykken and Auke Tellegen's twin studies (Lykken and Tellegen 1996).

In the twin studies by Lykken and Tellegen Well-Being measures of monozygotic and dizygotic twins were taken and compared after 10 years. Twin A's Well-Being score was compared with twin B's Well-Being score 10 years later. While the cross-twin, cross-time correlation for dizygotic twins was almost zero, the authors concluded from the data of the monozygotic twins, that the *stable* component of subjective well-being could be heritable by up to

80%. Towards that result Diener argues that when measured over such a long period as 10 years the results are naturally more likely to be influenced by stable factors like the genes and less likely by events that might have influenced current well-being (Ed Diener, E. M. Suh, Lucas, et al. 1999, p. 279). It is therefore argued that the 80% as set by Lykken and Tellegen is probably set too high, hence Lyubomirsky and Sheldon simply work with a 50%

estimate, arguing others also share this estimate (see Lyubomirsky, K. M. Sheldon, and Schkade 2005, p. 113). Another interesting study by Sonja Lyubomirsky to mention here is "Personality Fit And Positive Interventions" (Lyubomirsky, Tkach, and K. M. Sheldon 2004; Schueller 2012). The SWLS and PANAS were used to measure life satisfaction and PA, together with the "Big Five Inventory" to distinguish introverts and extraverts. The study was successful in establishing that interventions differ in effectiveness between introverts and extroverts. While writing a thank you letter and reading it to the person it is addressed to aloud proved to have a very high effect on extroverts, introverts were better served by reminding themselves of three good things that had happened each day and how they were part of it.

On the basis of the presented literature on the pleasant life, to which Seligman refers often, I argue that Seligman is *mostly* interested in providing intervention.

#### **4.3.2 The Engaged Life**

The second component of Seligman's understanding of happiness is mostly based on research done by Mihalyi Csikszentmihalyi who articulated the "Flow"-theory. Through qualitative interviews and observations on very successful people like artists and promising students Csikszentmihalyi concluded, that people are happiest, when they are in the "flow" state, primarily driven by self-set goals (Csikszentmihalyi 1990). Csikszentmihalyi observed artists being indifferent to hunger, fatigue and discomfort as long as work on a painting was going well. Through structured interviews and later the ESM technique he came to the conclusion that there is a fragile balance between perceived challenges and skills: "If challenges begin to exceed skills, one first becomes vigilant and then anxious; if skills begin to exceed challenges, one first relaxes and then becomes bored" (Nakamura and Csikszentmihalyi 2002, p. 89). However, it is not simply a balancing of challenges and skills. Flow mostly occurs when the person in question sees greater opportunities for action than on an average day and has the skills to adequately engage the challenge. Experiencing Flow leads the person to persist and return to the activity because of the experiential rewards, which in time will lead to fostering of this skill. Flow is a concept that relies on the person to be intrinsically motivated in her work. These autotelic personalities are least happy and motivated in apathy, whereas nonautotelic people did not find apathy conditions (low challenge, low skill) aversive.

The second caveat to flow when trying to link it to 'happiness' is that it is an amoral concept. It is "possible for people to seek flow in activities that are neutral or destructive to the self and/or the culture" (*ibid.*, p. 101). Being in the flow, means pursuing engagement, where time passes quickly and one is completely absorbed by the activity, while the attention is purely focused on what one aims to do. "Flow is defined as a psychological state in which the person feels simultaneously cognitive efficient, motivated, and happy (Moneta and Csikszentmihalyi 1996, p. 227). In order to get data on mood, Csikszentmihalyi

devised his own questionnaire with 26 mood-related adjectives. The Experience Sampling Method was developed to be able to apply this questionnaire successfully. At random times of the day participants filled out the questionnaires. The mood questionnaire that Csikszentmihalyi used consisted of 26 adjectives related to mood, which were positioned on 13 opposite scales like "happy - sad", "ashamed - proud", "tense - relaxed" or "clear - confused". While the aim of the research on flow was to get data on the concentration, location, bodily state and people accompanying the study participants, the claim on how happy the participants were stems solely from one 7 point scale ranging from happy to sad (see Csikszentmihalyi and Reed Larson 1987, p. 535; and Moneta and Csikszentmihalyi 1996, p. 284). It is surprising how prominently Seligman refers to flow theory, since the authors themselves are quite unsure how the concept of flow is tied to 'happiness': "Does the sum of flow over time add up to a good and happy life? Or only under certain conditions, that is, if the person develops an autotelic personality and learns to enjoy high challenges?" (Nakamura and Csikszentmihalyi 2002, p. 102).

Still, the wish to enhance the state of flow seems obvious enough and Seligman proposes the enhancement of this status by identifying people's highest strengths and then help them to find opportunities to engage in situations with their respective skills. For this purpose, Seligman and contributors published *Character Strengths and Virtues* (Peterson and Martin E. P. Seligman 2004), a classification of positive traits that they designated as character strengths. *Character Strengths and Virtues* is supposed to fulfill the function of an anti Diagnostic and Statistical Manual of Mental Disorders 4th ed. (DSM-IV). A proposed character strength had to pass a total of 10 criteria in order to be accepted (see *ibid.*, pp. 17–27). A short selection: Criterion 1 states that a strength can be good for others as well, but it has to fulfill the individual in question. Criterion 2 states that strengths have to produce desirable outcomes and need to be morally valued. Criterion 5 states that a strength needs to manifest in a range of behaviors, have a degree of generality across situations and stability across time – so it can be assessed. The resulting 24 character strengths could be summarized under six moral virtues<sup>69</sup> that can be found "across different cultures and throughout time"<sup>70</sup> (*ibid.*, p. 28). *Character Strength and Virtues* then discusses character traits of the core moral virtues and how to measure them through operationalizations. The resulting instrument Values in Action Inventory of Strengths (VIA-IS) questionnaire is measuring the 24 character strengths related to the 6 core values (*ibid.*, p. 627).<sup>71</sup> The resulting VIA-IS score can then be used find activities that a specific person would feel engaged in.

---

<sup>69</sup>The six core moral virtues are: Wisdom and knowledge, courage, humanity, justice, temperance and transcendence. Each of these moral values had several character traits attached to them, i.e. creativity and open-mindedness to wisdom and knowledge, bravery and persistence to courage, love and social intelligence to humanity etc. (Peterson and Martin E. P. Seligman 2004, pp. 29–30). These virtues are later on called "the high six".

<sup>70</sup>These virtues are then compared to virtues of traditional lore, like 500 B.C.E Confucian virtues, Buddhist virtues, Jewish, Christian and Islamic virtues, and the Nicomachean Ethics

<sup>71</sup>The questionnaire can be found on [www.viastrengths.org](http://www.viastrengths.org).

### 4.3.3 The Meaningful Life

The third happy life involves the pursuit of meaning. One's signature strength in the meaningful life is used "to belong to and serve something that one believes is bigger than the self" (Martin E. P. Seligman, Rashid, and Parks 2006, p. 777). These ideas are potentially rooted in religion, politics, family, nation etc. The subjective sense of meaning is said to strongly correlated to happiness. While this claim intuitively makes sense, Seligman presents few studies that support the argument. Seligman refers to a metastudy by Sonja Lyubomirsky, Laura King and Ed Diener in which a total of 225 papers were analyzed. Only one of the papers is looking at the correlation of positive emotions and meaning at the job (Lyubomirsky, King, and Ed Diener 2005, p. 807) and within the study only one of a total of 43 questions concerns itself with meaning and is therefore a rather subordinate claim of the study (Staw, Sutton, and Pelled 1994, p. 67). Seligman goes on to claim that the greatest benefits of pursuing meaning are to be able to transform the perception of a person's circumstances to look more fortunate. However, the study that Seligman refers to concerning this claim compares two groups of people in how they tell their life story and notices one group having a higher likelihood to tell their story in a specific narrative (McAdams et al. 1997). The study was not linked to 'happiness' and the authors even state that the life stories of the members of one group "are no more optimistic in narrative tone and contain no greater number of positive [...] affect experiences, than those life stories" by the other group (ibid., p. 687).

## 4.4 Conclusion on The Three Happy Lives

In summary, Seligman's empirical basis for the three happy lives is at times rather thin. Too often studies were referenced that (sometimes explicitly) do not seem to have a link to the research on 'happiness', or there are inconsistencies in how he uses the terms. For the description of the engaged life Seligman refers to the Csikszentmihalyi's flow and his and Peterson's *Character Strengths and Virtues*. In *Character Strengths and Virtues* Seligman argues per criterion 2 that character strengths need to be morally valued (see Peterson and Martin E. P. Seligman 2004, p. 19), while Jeanne Nakamura and Mihaly Csikszentmihalyi explicitly acknowledge the amorality of flow (Nakamura and Csikszentmihalyi 2002, p. 101). Seligman then proposes to enhance flow by the character strengths which have to satisfy moral evaluation (Peterson and Martin E. P. Seligman 2004, p. 19). This might not necessarily be a problem but the fact that it is not mentioned generates the impression that Seligman is not so much interested in sketching a consistent theory of 'happiness' but rather erratically looking for any clue to support his classification of character strengths on which his therapy interventions are based. The three lives are presented as "scientifically manageable components" of 'happiness' (Martin E. P. Seligman, Rashid, and Parks 2006, p. 776). However, the only 'scientific' link between these three components is that the

absence of all of them is claimed to be a cause of depression.

## 4.5 Flourish

In 2011 Seligman came to the conclusion that his "authentic happiness" theory<sup>72</sup> (Martin E. P. Seligman 2002) is wrong. In his book *Flourish*, published in 2011 Seligman reports having changed his mind on what the research topic of positive psychology should be about:

"I used to think that the topic of positive psychology was happiness, that the gold standard for measuring happiness was life satisfaction, and that the goal of positive psychology was to increase life satisfaction. I now think that the topic of positive psychology is well-being, that the gold standard for measuring well-being is flourishing, and that the goal of positive psychology is to increase flourishing" (Martin E. P. Seligman 2011, p. 13).

One of the few scientific references in this book is on Felicia Hupperts and Timothy So's "Flourishing Across Europe" (Huppert and So 2013). Huppert and So write that several definitions and scales for flourishing exist, but that there is no agreement yet. In their own study they defined components of flourishing as "the mirror opposite of the symptoms of the common mental disorders" (*ibid.*, p. 841). Seligman's own suggestion is that the well-being theory's flourishing should be measured by the elements positive emotion, engagement, relationships, meaning and achievement – PERMA for short. In the book Seligman gives no consistent reasoning why these five elements make up the new well-being theory, except for intuition and personal anecdotal evidence. When I asked Martin Seligman about the empirical evidence behind the five elements of well-being, he responded: "[I]t was not an empirically based proposal. [I]t was a priori [sic!] lots of empirical data since to test it and its psychometrics".<sup>73</sup>

Well-being theory and its PERMA components were first introduced in *Flourish*, a popular book on 'happiness', littered with personal and historical anecdotes. Very few of the references made in the book are explicit or complete with author, publication and year. Every PERMA reference I have found in scientific publications refers to this very book. Seligman further writes:

"Well-being is a construct, and happiness is a thing. A »real thing« is a directly measurable entity. Such an entity can be »operationalized« - which means that a highly specific set of measures defines it. [...] Well-being theory denies that the topic of positive psychology is a real thing; rather the topic is a *construct* – well-being- which in turn has several measurable elements, each a real thing, each contributing to well-being, *but none defining well-being* (Martin E. P.

---

<sup>72</sup>Not discussed in this thesis.

<sup>73</sup>Personal communication per email, 25. October 2021.

Seligman 2011, pp. 14–15)

Well-being definitely is a construct, since it is not directly observable, but so is 'happiness'. What Seligman probably means is that well-being is a composite of many things while we accept 'happiness' to be 'one' thing. Then again, we all would probably agree, that there are many elements contributing to 'happiness'. One could argue that well-being is just a made up word summarizing many contributing phenomena, but the same could be argued about 'happiness'. 'Happiness' being a "real thing" in contrary to the construct of well-being is another interesting claim. The attribute "real" isn't somewhere to be found inbetween "construct" and "operationalization". I'm sure we all would agree, Seligman included, that temperature is real. Temperature is a construct. Constructs are not *directly* observable, which holds true for all three of the terms. According to Seligman the "real thing" 'happiness' is directly measurable by means of operationalization. Then again, "operationalization" is the very process of making unobservables observable. The "real thing" elements of well-being include "meaning" and "engagement" according to Seligman's PERMA theory. Are they directly observable? In the paragraph about meaning Seligman argues that meaning is not solely a subjective state. "The dispassionate and more objective judgment of history, logic and coherence can contradict a subjective judgment. Abraham Lincoln [...] judged his life to be meaningless, but we judge it pregnant with meaning" (Martin E. P. Seligman 2011, p. 17). In that strain Seligman argues, that Jean-Paul Sartres play *No Exit* (Huis clos) while judged meaningful by its author, "now seems wrongheaded" (*ibid.*). Maybe it is just that I disagree with Seligman on Huis Clos, but I do not see the "directly measurable" part of the "real thing" meaning. My alternative suggestion is to understand Seligman's proposal as a new proxy construct called "flourishing". 'Happiness' lacks a clear operationalization. Seligman now limits the possibilities of operationalization by suggesting that 'happiness' could empirically be approached from a flourishing perspective using the PERMA instruments. As such this is an interesting proposal given that the PERMA profiler allows a much broader assessment than the usual SWB instruments like the SWLS can provide.<sup>74</sup> That said, neither 'happiness' nor SWB nor flourishing are directly measurable. All three need to be operationalized in order to create instruments that render these constructs empirically tangible. However, while the big dispute about 'happiness' has not been resolved yet, the creators of SWB and flourishing did outline what the respective constructs should encompass. In the case of SWB, measures should include global measures of life satisfaction and no measures that we would deem "objective". In the case of flourishing, measures should include the elements lending the name to the PERMA profiler: Positive emotion, engagement, relationship, meaning and accomplishment. Since both of them refer to

---

<sup>74</sup>In comparison to the 5 questions of the SWLS that only concern the satisfaction with one's life, the PERMA profiler encompasses 23 questions, among others asking the respondent about her physical health, to what extent she receives help and support, how lonely she feels during her daily life and to what extent she feels loved (see Ed Diener, Emmons, et al. 1985; and Butler and Kern 2016).



'happiness', while being closer to operationalization, I categorize them both as proxy constructs.

## Chapter 5

# Methodological Discussion and Criticism of Positive Psychology

I started to work on the topic of positive psychology with the critical literature in mind (foremost Kleiner 2015; Ehrenreich 2009; Taylor 2001; Bayertz 2010, 2012; Lazarus 2015). The picture I painted in my head was one of a sub-discipline of psychology whose scientific motivation was put into doubt by the underlying neocapitalistic political drive and questionable methodological work that had turned into a cash machine by selling therapy to people wanting to maximize their 'happiness' and feeding popular science literature.<sup>75</sup> Eva Illouz and Edgar Cabanas summarize:

"It seems clear, that positive psychology could not have thrived on the basis of its science alone. The field is characterized by its popularity as much as its intellectual deficits and scientific underachievement. [...] It could be argued that positive psychology is little more than ideology recycled in the form of charts, tables and number-filled diagrams; an easily marketable pop psychology touted by scientists in white coats" (Cabanas and Illouz 2019, pp. 32–33).

Many a times a scientific conclusion like 80% of stable happiness being hereditary (see Lykken and Tellegen 1996) was properly referenced and explained in a paper. But the reason why this number should be reduced to 50% (see Lyubomirsky, K. M. Sheldon, and Schkade 2005, p. 113) was harder to come by. The papers using the 50% number did not explain how the number came to be, neither did the papers that were referenced in order to argue for this number. It took a lot of digging in all given references until I could conclude with confidence that the "50%" was not a calculated number but a simple estimate based on the conviction that 80% 'must be too high', since its computation only tracked stable

---

<sup>75</sup>One lead that I would have wanted to follow were the interconnections of the [John Templeton Fund](#), its potential political creeds as manifested by the funding areas "Philosophy & Theology" and "Individual Freedoms & Free Markets" and positive psychology. Martin Seligman's initiative attracted the attention of self-made billionaire [Sir John Templeton](#) who initially funded the projects through [research grants and prize money](#). Seligman became the principal investigator of several Templeton Religion Trust and John Templeton Foundation research projects and now has a chair in the steering committee and serves as advisor to the John Templeton Foundation according to his [CV](#).

”Well-Being” of monozygotic twins.<sup>76</sup>

Then there are methodological inconsistencies like Barbara Fredrickson and Marcial Losadas ”Complex Dynamics of Human Flourishing” (see Fredrickson and Losada 2005a). Based on Barbara Fredrickson’s ”broaden-and-build-model”, the team argued that business teams that expressed higher positivity in their meetings must perform better.

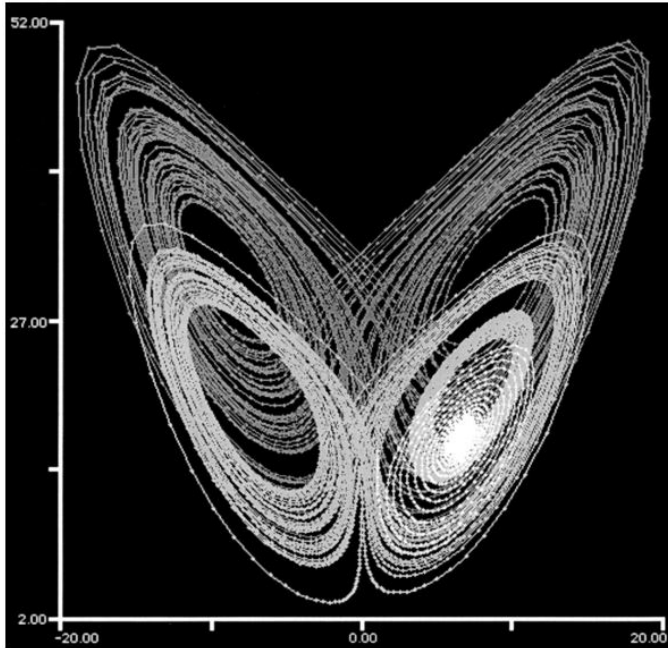


Figure 5.1: Representation of the dynamic structures describing high-performance, medium-performance, and low-performance teams (ibid., p. 682).

The data from the conducted study observing 60 business teams confirmed the expectations. Fredrickson and Losada calculated that teams with a measured positivity ratio at or above 2.9013 times more positive than negative things discussed were the most high-performant and mentally flourishing.

The authors then purported a mathematical model based on the ”Lorenz system” used in fluid dynamics representing the performance and PA/NA ratios in emotional space as seen in figure 5.1. Nicholas Brown and Alan Sokal (the latter known in the social science through the science wars) re-

futed the mathematical model in ”The Complex Dynamics of Wishful Thinking”, calling it an ”abuse of mathematics” (Brown, Sokal, and Friedman 2015, p. 30). Fredrickson has since withdrawn the mathematical model elements, including the positivity ratio of 2.9013 (see Fredrickson and Losada 2013).

It is understandable that happiness research (the broader historical field) and positive psychology specifically (since 1998) has come under attack, also since they do not consider data from the social sciences. This picture of a dubious science is not entirely wrong. But

<sup>76</sup>The paper by Lykken and Tellegen (1996) is a study based on the data of the ”Minnesota Study of Twins Reared Apart” (MISTRA) (For more information on this very interesting study please see Segal 2012). The paper by David Lykken and Auke Tellegen compared ”Well-Being” measures (not *Subjective* Well-Being, but a different proxy construct) of monozygotic and dizygotic twins reared apart in intervals of up to 10 years. They argue, that the correlation of a Well-Being measure of a dizygotic twin with the measure of it’s sibling 10 years later was ”essentially zero” (Lykken and Tellegen 1996, p. 188). The same cross-twin comparison of two monozygotic twins over 10 years equaled to 80% however. The above mentioned paper is often used to argue for a certain heredity of ’happiness’, while applying the results of a study measuring ”well-being” to studies using ”subjective well-being”. For my criticism on why the results of studies using different proxy constructs should only be applied cautiously to one another please see page 16 in this thesis. For criticism on MISTRA and why it does not have the final word on the ”nature versus nurture” discussion please see Joseph (2022).

neither is it complete. In the defence of happiness research one could argue that they very well did consider demographic factors, in the 1960s and 1970s, and insofar as contemporary happiness research portrays itself in line with social indicators research they could argue that socio-demographic factors play a very subordinate role in study participants perceived happiness. Therefore, to some degree the criticism can be refuted on the basis of historical research done.

While I certainly agree with Eva Illouz's and Edgar Cabanas' assessment that positive psychology plays neatly into a neo-liberal agenda (see Cabanas and Illouz 2019), I would like to spend the next few paragraphs analyzing how science is conducted and argue that positive psychology, while certainly having its flaws, is not faring that poorly after all.

### 5.1 On Comparability and Scientific Progress

Hasok Chang, in his great book on water argues that water, an element as understood in the phlogiston theory, became H<sub>2</sub>O over the course of time during the chemical revolution (see Chang 2014, 133pp.). Scientists often arrive at what Chang calls the "problem of nomic measurement".<sup>77</sup> In order to break the cycle between wanting to develop a theory but needing a specific measure in order to go forward and wanting to develop an instrument but needing a theory in order to conclude what to measure, many methods seem permissible. During the transition of water from an element to a chemical formula of atoms not every argument was based on 'good methodology'. John Dalton, claiming water is HO in 1808 simply applied his "rule of greatest simplicity". Through eletrolysis Dalton only knew of one compound of hydrogen and oxygen each. "Why should he have assumed that the composition of water was not the simplest possible, namely a one-to-one combination of hydrogen and oxygen atoms" (ibid., p. 138)? While Dalton's "rule of simplicity" convinced most of his peers for practical reasons, the doctrine proved to be insufficient with more complex molecular formulas. By operationalizing the relative volumes of atoms in chemical reactions, Wilhelm Hofmann was able to solve the mystery in 1865, and suggested that water is hydrogen and oxygen combined in a 2:1 volume ratio.

Uljana Feest argues that (operationally defined) concepts can be used as research tools that help to temporarily fixate conditions of application (see Feest 2010, p. 179). In that vein, Dalton's "rule of simplicity" allowed an operationalization of atomic weight. While the assumption retrospectively had to be corrected, it proved to be a milestone towards the Formula of H<sub>2</sub>O.

"Here is a guiding idea to help us understand the spirit of nineteenth-century atomic chemistry: in order to learn something real about atoms, *do* something with them. [...] For *physical* atoms, operationalization mostly involved some advanced technology or high-powered statistical reasoning, neither of which

---

<sup>77</sup>In this thesis shortly discussed on page 30.

became available until the late nineteenth century. For *chemical* atoms, operationalization came much sooner and much more easily. By operationalizing the concept of atoms, chemists learned to do a pragmatic sort of metaphysics” (Chang 2014, p. 153).

Bradburn operationally defined happiness as having more PA over NA and since he gave no detailed account how the questionnaire items for PA and NA were constructed, it must be assumed that it must have been based on intuition alone (see Bradburn and Caplovitz 1965). The resulting two clusters did not correlate, to Bradburn’s surprise. The clusters allowed further research by Andrews and Withey confirming the independence of PA and NA and lead to the third factor of ”cognitive evaluation” being discovered (see Andrews and Withey 1976). This discovery allowed Ed Diener to converge on his three markstones of subjective well-being (see Ed Diener 1984) and finally to dismiss the independence of PA and NA as a statistical artifact produced by memory bias and too long time frames. Bradburn’s conclusion on the independence of NA and PA can retrospectively be considered ’wrong’, but he definitely played his part in the research on happiness so it could arrive where it by now did. While these operationalizations sometimes seem overly crude, I agree with Chang’s sentiment that ”operationalization is an achievement” (Chang 2004, p. 86).

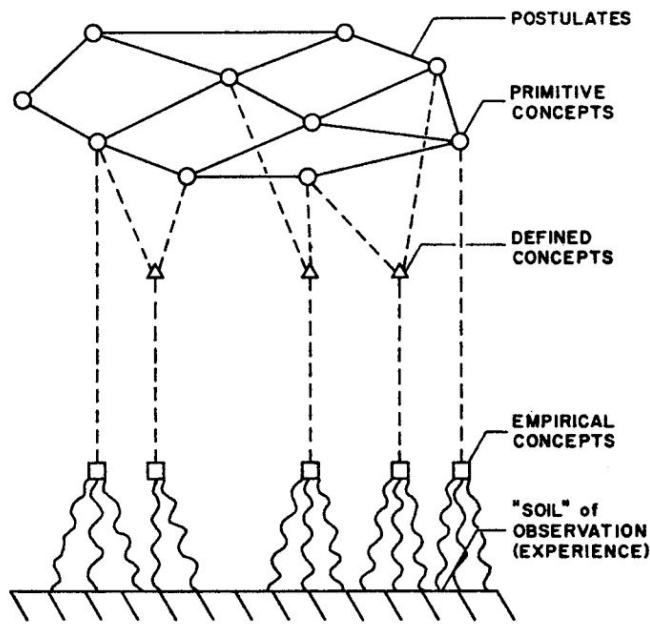


Figure 5.2: Feigl’s representation the connection between postulates and observations (*ibid.*, p. 6).

research teams to join, since the conditions to work on something comparable are clear.

I concede that proxy constructs are probably on the same level as Herbert Feigl’s ”primitive concepts” as seen in figure 5.1. No definite measure has been predetermined yet, but SWB is already based on some empirical evidence and less broad than ’happiness’, the

In this thesis I have argued for the separation of constructs and proxy constructs for three reasons.

- (1) Constructs are on a more abstract level. A proxy construct refers to one construct, but any construct can be referred to by many proxy constructs.
- (2) Confusing the two might lead to vague science since two proxy constructs might not be comparable even though they refer to the same construct.
- (3) As in the example of SWB, being transparent about the semi-operationalized character of a proxy construct might create a new sub-field and enable other re-

superordinate notion. Out of subjective well-being followed the Satisfaction With Life Scale.

Is subjective well-being as measured by the Satisfaction With Life Scale 'happiness'?

No!

It can't be.

'Happiness', as an unoperationalized construct does not - and cannot - correspond to anything in the physical realm.<sup>78</sup> Every construct needs to be operationalized in order to be measured and it can be operationalized in different co-existing ways.

Percy Bridgman's operationalism was a reaction to very different ways of measurement like measuring with a meter stick and specifying distance in light years sharing the same word in common tongue (see Bridgman 1927). Operationalization is the answer to non-homogeneous concepts.

"In practice, however, scientists do not recognize multiple concepts of length, and Bridgman was willing to concede that it is allowable to use the same name to represent a series of concepts, if the different measurement operations give mutually consistent numerical results in the areas of overlap" (Chang 2014, p. 145).

Operationalizations have a vertical and a temporal quality to them. As seen in Feigl's illustration in figure Nr. 5.1, postulates on the vertical axis get more and more explicit, until an operationalization allows an actual measurement. Every operationalization is also 'merely' a suggestion for a measure that might be linked to an unobservable construct. Only by comparing the operationalization to other operationalizations over the course of time will answer the question whether it is a good measurement for the intended use. A measurement possesses value in term of that specific operationalization of the postulate. But said measurement was only possible due to the operationalizations on the vertical axis. The measure does not directly relate to the postulate, nor can we independently judge whether the operationalization was in spirit of the higher level concept.

In this thesis I have used the term "proxy construct" to describe the epistemic status of SWB. With reference to Chang (2004) and Chang (2014), I argue with this further case study on 'happiness' for the importance of operationalizations that allow scientists to *do* something with their purported epistemic thing. I argue that proxy constructs are an important step on the way for operationalizations to emerge. Additionally, proxy constructs facilitate the comparison of different operationalizations, since two operationalizations based on the same proxy construct share the same base assumptions upon which the proxy construct is defined. As such, conclusions drawn from an operationalization of a proxy construct from one team are easier to integrate into the experimental system of another team provided they share the same base assumption by doing research on the same proxy construct. Within the literature on happiness, I found SWB to be the most

---

<sup>78</sup>See Chang (2014, 207ff.) for an elaborate assessment on constructs and operationalizations.

distinguishable proxy construct, with its pre-operationalization explicitly described as its "three hallmarks".<sup>79</sup> I argue that the transparency of what the proxy construct subjective well-being entails constitutes an important cornerstone of its success and facilitated other research groups to contribute operationalizations according to their own flavour to this proxy construct.

## 5.2 On Validation

The validation of an operationalization is making sure that what is measured equals to what is supposed to be measured and is called "construct validation". Denny Borsboom argues that construct validation as it is usually being done, is actually test validation, since not the constructs, but the tests are validated (see Borsboom 2003, 128 ff.). As shown on the illustration by Andrews and Withey in figure Nr. ?? on page ??, construct validation compares the overlap of the data of two instruments, being two operationalizations of a construct.<sup>80</sup> Borsboom argues that construct validity as it is being practiced, therefore does not test whether a certain operationalization is measuring the abstract construct, but how well such an operationalization is backed by evidence. In this regard I would go even further and claim that construct validity is merely showing the overlap of two operationalizations. A high overlap of two operationalizations is an indication that these operationalizations measure the same epistemic thing since the measurement scores seem to be somewhat comparable. I believe Borsboom holds scientists to too high standards when he argues that "[w]hen claiming that a test is valid, one is taking the ontological position that the attribute being measured exists and affects the outcome of the measurement procedure" (*ibid.*, p. 131). As argued above, in my opinion no argument about the actual existence of anything has been made. Only that said operationalizations seem to measure similar or the same epistemic thing, given the scores are similar. Borsboom criticises psychology for not having "unambiguous connections between »theoretical sentences« and »observational sentences«" (Borsboom et al. 2009, p. 137) in a positivist understanding. Further he disagrees with current psychologists' practice of substituting nomological networks (which according to Borsboom do not exist in contemporary psychology) with sets of correlations. The term "construct", the way it is used in contemporary psychology is used to refer to the theoretical term in a theory as well as the possible referent of the term (the epistemic thing 'happiness'). For the former I have proposed the term "proxy construct", if the description of the term partly defines how it is supposed to be operationalized. The tip of the iceberg invalidating construct validation as envisioned by Cronbach and Meehl (1955) in the eyes of Borsboom is that the authors' understanding of construct

<sup>79</sup>The three hallmarks can be found on page 46 in this thesis.

<sup>80</sup>In construct validation usually but not necessarily operationalizations of the same construct are used. Conversely, in the construct validation of the Satisfaction With Life Scale Diener also compared SWLS scores with Bradburn's PA and NA measures. Since the SWLS is not supposed to tap into affect, but only into life satisfaction, low correlation between SWLS and PA/NA scores were seen as a positive sign for construct validation of the SWLS (see Ed Diener, Emmons, et al. 1985, p. 73).

validation include *interpretations* of test scores, resulting in validity being the outcome of test score interpretations rather than of tests. According to Borsboom however, real construct validation should not refer to some test scores (based on operationalizations) but check whether a certain test is measuring what it is supposed to measure. He argues for a realist approach to measurement and for a kind of validity that does not refer to test score interpretations, but to the properties of the tests. In his understanding of construct validity, Borsboom renames it *test validity*, which "is not a function of evidence, but a function of truth" (Borsboom et al. 2009, p. 138). Instead of construct validity as it is being practiced today, Borsboom et al. suggests test validity should "explain how the psychological attribute acts to cause variation in [...] measurement outcomes" (*ibid.*, p. 165). In my view scientists then are back at the "problem of nomic measurement" that is haunting all theory-laden observations. Further I can only refer to Hasok Chang's lengthy arguments that the success-truth link in science is highly suspect and I encourage the reader to consider the respective passages in Chang (2014, 227ff.) and (*ibid.*, 240ff.). I understand the philosophical desire to have such a kind of validation that checks the link between an abstract construct and a concrete operationalization. However, I think such a point of view neglects the actual scientific practice of breaking the chain of 'the problem of nomic measurement' through operationalizations. As already mentioned by Angner (2013) I would argue that some happiness scholars fall into the same trap through trait reification where scores of tests are mixed up with the abstract construct 'happiness'.

### 5.3 In Defence of Temporary Claims

Above I have argued that construct validation does not actually validate the construct, but only that operationalizations that probably tap into a similar phenomenon. I agree that it is good evidence for a stable referent when instruments of several proxy constructs pass construct validation. My main complaint is that the term "construct validation" promises a more qualitative assessment of the construct than this statistical tool can actually provide. As Borsboom rightly notes, the arguments for the validity come from the tests and not from the description of the construct. Above I have also argued that operationalizations have a temporal quality to them. I wholeheartedly agree with Uljana Feest that by using operational definitions "scientists were *partially and temporarily* specifying their usage of certain concepts by saying which kinds of empirical indicators they took to be *indicative of* the referents of the concepts" (Feest 2005, p. 133). Scientists trapped by the 'problem of nomic measurement' might define a concept operationally because it allows them to measure something. Such an operationalization need not be correct, as the example by Hasok Chang with the water formula of "HO" shows. Nevertheless, in the long run, through a process he calls "epistemic iteration"<sup>81</sup> flaws in a current operationalization will be found

<sup>81</sup>A process in which successive stages of knowledge, each building on the preceding one, are created in order to enhance the achievement of certain epistemic goals. It differs crucially from mathematical iteration



and another one agreed on that allows scientists to *do* further research. In that vein, Uljana Feest argues that the replication crisis in psychology tells us less about the quality of research but about the focus of that discipline on exploration (see Feest 2019).

#### 5.4 Assessing Positive Psychology

In chapter 2 I have shown how Ed Diener worked on operationalizations of subjective well-being by setting up multi-staged studies. He introduced new instruments, which supposedly serve the purpose of proxy construct better, but need to be validated against the old instruments first. This is a nice fit with Rheinberger's description of technological objects claiming that they are usually shaped by "old tools". These serve as well-needed points of reference when new tools are being developed. If the new tools measure the same construct and are 'better', by *i.e.* being more reliable in which factors they pick up or being faster to apply, they can replace the old instruments in the long run. "Yet in the long run, a technical system [object] may become completely replaced by systems which embody the current, stabilized knowledge in a more efficient way" (Rheinberger 1992b, p. 323). Bradburn's PA and NA Scale serve as examples for such replaced tools. Since Bradburn gave no indication why the tools were constructed in this way, positive affect and negative affect are now mainly measured by applying the Positive and Negative Affect Schedule (PANAS), a 20 item questionnaire which was constructed through factor analysis of 60 emotional terms (D. Watson, Clark, and Tellegen 1988).

This however, brought its own set of problems, discussed by Alexandrova and Haybron (2016, p. 1105). Alexandrova and Haybron make the point that factor analysis allows happiness scholars to avoid having to settle on a theory of 'happiness' (and/or a proxy construct) in order to choose questionnaire items and instead use solely statistical arguments for tuning their instruments. While agreeing, I also think that arguing for the items of an instrument with factor analysis is still better than having no explicit arguments how the instrument items were constructed at all.

Ed Diener's multi-staged studies are built consistently within and newer studies are based on the results of previous ones. This structure fits the description of "epistemic iteration" by Chang pretty well. Martin Seligman has a different approach to scientific work. As argued, the PERMA theory was published without prior empirical data. As PERMA has since been tested and successfully applied, we could portray the 'premature' publication of PERMA as a further example of 'breaking the chain of nomic measurement'.

Ed Diener's approach of doing science and publishing results is definitely easier to follow as a reader. Both ways of working seem to be permissible however. As far as Ed Diener and Martin Seligman can be portrayed representative for the field of subjective well-being and positive psychology they seem to follow different aims. Ed Diener unveiled the fine-

---

in that the latter is used to approach a correct answer that is known, or at least in principle knowable, by other means" (Chang 2004, p. 253).

grained mechanism of life satisfaction, opposite character traits and uncovered previous instruments producing statistical artifacts by not differentiating between intensity and frequency. Martin Seligman as a clinical psychologist is rather interested in developing and therapies and interventions.

A therapy based on his "three lives"<sup>82</sup>, which was informed by his work on the 'anti-DSM' *Character Strengths and Virtues* was able to reduce depressive symptoms of mildly and severely depressed patients significantly better than conventional therapy with and without the application of drugs. Further, Seligman's therapy also raised life satisfaction of the patients as measured by the SWLS significantly compared to conventional therapy (see Martin E. P. Seligman, Rashid, and Parks 2006). Ed Diener was rather interested in uncovering an 'inward lead', subjective operationalization of 'happiness' while developing instruments to do so. Martin Seligman used the uncovered knowledge, worked on therapies and then applied the SWB in order to prove their superiority to existing ones.

Examples of unsound work can be found within positive psychology. That however does not mean that the whole sub-discipline is doomed to fail. As I have argued in this chapter, not only are temporal assumptions, operational definitions and construct validation as used by positive psychology permissible in other sciences, they have also led to successful scientific projects in the past. In that sense, whether positive psychology will ever come up with an unilaterally accepted operationalization of 'happiness' that is good enough<sup>83</sup> it is too early to say. As far as scientific progress can be analyzed however, positive psychology has an impressive track record.

---

<sup>82</sup>Described on page 76.

<sup>83</sup>Hasok Chang on water: "I conclude that water is H<sub>2</sub>O, but also other things, *really*. [...] But we can perhaps accept these [further conditions that need to be fulfilled] as well-understood and thoroughly agreed »fine print« that does not change the main message, to that »Water is H<sub>2</sub>O« appears as an unproblematic approximation to the truth" (Chang 2014, pp. 203-204).



## Chapter 6

### Conclusion

This thesis attempted to evaluate whether the approaches used by philosophers of science to assess scientific findings can be used for the social sciences as well. In particular, I looked at the attempts to measure 'happiness' by positive psychology and its predecessors. To this aim, several approaches that have been used in assessing scientific findings are introduced and adapted to the psychological studies portrayed in this thesis.

To the best of my knowledge, a history of 'happiness' research in general and of positive psychology specifically does not exist yet.<sup>84</sup> In the writings of contemporary scholars of happiness, only anecdotal stories can be found. While a comprehensive history of the sub-discipline is still missing, this thesis aimed at reconstructing some of the central scientific arguments and instruments used in the psychological science of happiness during the last decades. This experimental history of happiness research follows the journey of a philosophical understanding of 'happiness' to an empirical research topic. While being researched from a broader social science perspective first, measuring happiness became a psychological issue mostly.<sup>85</sup>

The phenomenon 'happiness' cannot be directly observed and has to be treated as a mental construct. Theories of the topic 'home in' on what parameters of human experience lead people to be happy. Operationalizations then define which specific attributes are measured with which specific instrument and refer back to the original construct of 'happiness' as one possible way of rendering the construct measurable. In this field exist many semi-operationalized constructs, that define a certain approach to what kind of phenomenon the authors expect 'happiness' to be. I suggested the term "proxy construct" for these kinds of constructs that are still not observable directly, but already pre-define what kind of attributes should be measured. In accord with Angner (2010) I argue that many scholars of positive psychology are guilty of score reification, expecting their instruments to perfectly

---

<sup>84</sup>McMahon's *Happiness* (McMahon 2006) is a very interesting read for anyone interested in the topic, but focuses on the change of philosophical interpretation of 'happiness' since the beginning of recorded history.

<sup>85</sup>There are scientific publications from other disciplines on the subject, *i.e.* economics come to mind. But the actual measure of 'happiness' is usually done via psychological instruments, paired with economic data and theorizing (see for example Frey and Stutzer 2002b).

represent the measured attribute and of trait reification, confusing the operationalization of their proxy construct with 'happiness', the nonobservable construct. I defended the position that since 'happiness' is not measurable per se and partly contradicting operationalizations of this construct are feasible, operationalizations and proxy constructs should always be referred to by their name and never summarized within the superordinate construct 'happiness'.

In this thesis I mostly focused on the work of Ed Diener and Martin Seligman regarding happiness research. Both are towering figures in the field, touted "the father of modern happiness research" and "the father of positive psychology" respectively by their peers. Diener created the field of subjective well-being (SWB) research by defining this proxy construct in an open fashion in such a way that many previous studies could already be counted towards SWB-research. Martin Seligman founded the movement of positive psychology, becoming it's figurehead, while working on interventions leading people to become more happy. Other individual researchers or research groups would have served the purpose of portraying happiness research just as well. I chose these two individuals because they very much differ in research goal and style, while at the same time complementing each other. Therefore, with a loose interpretation of Chang's "active (scientific) realism" (Chang 2014, 203ff.) we could argue that "anything goes" (Feyerabend 1982, p. 28) as long as it leads to an operationalization.

In the first chapter several approaches and vocabulary on assessing a science are introduced. In chapter two the history of an empirical debate about 'the good life', emerging after the middle of the 20th century is explained.

Chapter three studies the scientific process of dealing with the 'problem of nomic measurement' in the research group of Ed Diener. Empirical research usually struggles with theory needing empirical data, while instrument creation relies on a sophisticated theory, specifying what to measure. Diener and his research group were pioneers in developing new instruments and strategies to reject alternative explanations. New instruments had to be validated with old instruments in order to prove that they measure the same "epistemic thing" (Rheinberger 1992a,b), while at the same time producing data that is different enough so new scientific claims can be based on them. While Diener follows his scientific findings, it is interesting to note that his definition of 'happiness' is based on two minor factors, negative affect (NA) and positive affect (PA). The reasoning for the choice of factors was not based on empirical explanatory force, but on the fact that these two minor factors were better understood and easier to differentiate from other factors. The main factor "life satisfaction", while empirically explaining more of the variance in measurement scores is not part of his definition of 'happiness'.

In chapter four I investigate the work from Martin Seligman. I come to the conclusion that his work deviates from the aforementioned work by Ed Diener in goal and 'scientific style'. Coming from clinical psychology, Seligman's aim was to create interventions to

boost one's own happiness, while relying on non-empirical assumptions for the creation of these interventions and therapies. Nonetheless, Seligman's therapies were successful in ameliorating the suffering of clinically depressed patients when measured with Diener's instruments and his own.

In chapter five I discuss the philosophical aspects of doing science on unobservable constructs and some of the criticism leveled against positive psychology. While initially sharing the criticism, I eventually defend positive psychology against most of its critics. I argue that positive psychology, when only looking at the scientific side of it and compared with other historical cases, has had its fair share of shortcomings and missteps, but is still producing a meaningful outcome. It is too early to conclude on this research topic as a whole, since it may well lead to an all-encompassing definition of happiness in the years to come.

While I first dismissed the question of what 'happiness' as understood by positive psychology *is* in the introduction, I would like to ponder over that question for a bit at the end. 'Happiness', as psychologically understood<sup>86</sup>, are basic assumptions about its symptoms, measured where observable and validated by the overlapping of empirical data of several of these assumptions. A selection of these assumptions are: people that are happy make their tick at the upper end of a scale in a questionnaire asking how happy they are, they have a better immune system and therefore go to the doctor less often, they are more successful at their jobs, happy people live longer, 'happiness' is a mental state, 'happiness' is a character trait, 'happiness' is hereditary, 'happiness' can be fostered by supporting the strongest character strength of an individual etc. Some of these assumptions seems contradictive, which hints to the point that positive psychology has not converged on one operationalization of 'happiness' yet. Therefore, in order to answer the question of what 'happiness' *is*, we have to decide on an operationalization. If we take Diener's take on 'happiness', as an answer, then 'happiness' is having more positive experiences per day than negative ones, while explicitly not caring for the intensity of these experiences.

Is this understanding of 'happiness' *good enough*, or is it, as Anna Alexandrova asks, only lowering the bar for the purpose of making it measurable? (see Alexandrova 2017b, p. 135) I would argue that 'happiness' as understood by positive psychology, although losing some of the 'flavour' of a philosophical understanding, gains unambiguity, an important attribute when having application in mind. Observers of the empirical research on 'happiness' have to keep in mind that the interpretations of what 'happiness' is, is still very broad and depending on the context different proxy constructs are being used. Further, the experimental systems of positive psychology can only handle an 'artificial' understanding of 'happiness', 'happiness' in an all-encompassing understanding is too impure to be handled by the available instruments.

However, from a historical perspective, many such endeavours have born fruit in the long

---

<sup>86</sup>Ergo the measured operationalizations of it's proxy constructs.

run, as seen in the works of Chang (2004, 2014). Out of a theoretical assumption an operationalization was created, which in turn generated a further theoretical assumption, whose operationalization informed and potentially corrected the first assumption. Therefore, while many of the assumptions about 'happiness' may be incomplete or wrong, the project as a whole can still be successful. Anna Alexandrova has made a very similar argument, hence I close with her words:

"[I]t is no good to decide ahead of time from a philosopher's pedestal what well-being<sup>87</sup> is and then declare that no measure can do justice to this notion. [...] If well-being is to play a useful role in life of today's industrial bureaucracies which live by numbers, it may have to be *made* measurable even if it was not initially (Alexandrova 2017b, p. 135).

---

<sup>87</sup>The term "well-being" in the quote has been terminologically defined as 'happiness' in this thesis.





## Bibliography

- Abramson, Lyn Y., Martin E. Seligman, and John D. Teasdale (1978). “Learned helplessness in humans: Critique and reformulation.” In: *Journal of Abnormal Psychology* 87.1, pp. 49–74. ISSN: 1939-1846, 0021-843X. DOI: [10.1037/0021-843X.87.1.49](https://doi.org/10.1037/0021-843X.87.1.49).
- Alexandrova, Anna (Dec. 2012). “Well-Being as an Object of Science”. In: *Philosophy of Science* 79.5, pp. 678–689. ISSN: 0031-8248, 1539-767X. DOI: [10.1086/667870](https://doi.org/10.1086/667870).
- (Aug. 24, 2017a). *Is There a Single Concept of Well-Being?* Vol. 1. Oxford University Press. DOI: [10.1093/oso/9780199300518.003.0001](https://doi.org/10.1093/oso/9780199300518.003.0001).
- (2017b). “Is Well-being Measurable After All?” In: *Public Health Ethics* 10.2, pp. 129–137. ISSN: 1754-9973, 1754-9981. DOI: [10.1093/phe/phw015](https://doi.org/10.1093/phe/phw015).
- (June 1, 2018). “Can the Science of Well-Being Be Objective?” In: *The British Journal for the Philosophy of Science* 69.2, pp. 421–445. ISSN: 0007-0882, 1464-3537. DOI: [10.1093/bjps/axw027](https://doi.org/10.1093/bjps/axw027).
- Alexandrova, Anna and Daniel M. Haybron (Dec. 2016). “Is Construct Validation Valid?” In: *Philosophy of Science* 83.5, pp. 1098–1109. ISSN: 0031-8248, 1539-767X. DOI: [10.1086/687941](https://doi.org/10.1086/687941).
- American Psychiatric Association, ed. (1994). *Diagnostic and statistical manual of mental disorders: DSM-IV*. Washington, DC. ISBN: 978-0-89042-062-1 978-0-89042-061-4.
- Andrews, Frank M. and Stephen B. Withey (1976). *Social Indicators of Well-being: Americans’ Perceptions of Life Quality*. New York and London: Plenum Press.
- Angner, Erik (2005). “Subjective Measures of Well-Being - a Philosophical Examination”. PhD thesis. Pittsburgh: University of Pittsburgh.
- (June 2010). “Subjective well-being”. In: *The Journal of Socio-Economics* 39.3, pp. 361–368. ISSN: 10535357. DOI: [10.1016/j.socec.2009.12.001](https://doi.org/10.1016/j.socec.2009.12.001).
- (May 2013). “Is it possible to measure happiness?: The argument from measurability”. In: *European Journal for Philosophy of Science* 3.2, pp. 221–240. ISSN: 1879-4912, 1879-4920. DOI: [10.1007/s13194-013-0065-2](https://doi.org/10.1007/s13194-013-0065-2).
- Aristotle (2018). *Nicomachean Ethics*. Ed. by F. H. Peters. Standard Books.
- Baars, Bernard J (1986). *The cognitive revolution in psychology*. OCLC: 469489555. New York: Guilford Press. ISBN: 978-0-89862-656-8 978-0-89862-912-5.
- Bakan, David (1966). “Behaviorism and American urbanization”. In: *Journal of the History of the Behavioral Sciences* 2.1, pp. 5–25.

- Balatsky, Galina and Ed Diener (1993). “Subjective Well-Being among russian students”. In: *Social Indicators Research* 28, pp. 225–243.
- Barschak, Erna (Oct. 1951). “A Study of Happiness and Unhappiness in the Childhood and Adolescence of Girls in Different Cultures”. In: *The Journal of Psychology* 32.2, pp. 173–215. ISSN: 0022-3980, 1940-1019. DOI: [10.1080/00223980.1951.9916089](https://doi.org/10.1080/00223980.1951.9916089).
- Bauer, Raymond A., ed. (1966). *Social indicators*. [Nachdr. des] 4. print., Cambridge, Mass. 1972. Technology, space, and society. OCLC: 257187331. Niedernberg: Repro Pfeffer. ISBN: 978-0-262-02019-0 978-0-262-52007-2.
- Bayertz, Kurt (2010). “Eine Wissenschaft vom Glück. Erster Teil: Was ist Glück?” In: *Preprints of the Centre for Advanced Study in Bioethics* 4, pp. 1–17.
- (2012). “Empirische Antworten auf philosophische Fragen? Zum Verhältnis von philosophischer Ethik und empirischer Glücksforschung”. In: *Preprints and Working Papers of the Centre for Advanced Study in Bioethics* 49, pp. 1–11.
- Beckham, Albert S. (July 1929). “Is the negro happy?” In: *The Journal of Abnormal and Social Psychology* 24.2, pp. 186–190. ISSN: 0096-851X. DOI: [10.1037/h0072938](https://doi.org/10.1037/h0072938).
- Binning, John F. (Feb. 22, 2016). *Construct*. In: *Encyclopædia Britannica*. Encyclopædia Britannica.
- Borsboom, Denny (2003). “Conceptual issues in psychological measurement”. PhD thesis. Amsterdam: Amsterdam University.
- (Sept. 2006). “The attack of the psychometricians”. In: *Psychometrika* 71.3, pp. 425–440. ISSN: 0033-3123, 1860-0980. DOI: [10.1007/s11336-006-1447-6](https://doi.org/10.1007/s11336-006-1447-6).
- Borsboom, Denny et al. (2009). “The end of construct validity”. In: *The concept of validity: Revisions, new directions, and applications*. Ed. by R. W Lissitz. IAP Information Age Publishing, pp. 135–170.
- Bradburn, Norman M. (1969). *The Structure of Psychological Well-Being*. Chicago: Aldine Publishing Company.
- Bradburn, Norman M. and David Caplovitz (1965). *Reports on Happiness*. Chicago: Aldine Publishing Company.
- Bräuer, Sebastian (Apr. 13, 2014). “Glücksforschung. Geld ist gar nicht so wichtig”. In: *NZZ am Sonntag*, p. 33.
- Bridgman, Percy Williams (1927). *The Logic of Modern Physics*. New York: The MacMillian Company.
- Brown, Nicholas J. L., Alan D. Sokal, and Harris L. Friedman (2015). “The Complex Dynamics of Wishful Thinking: The Critical Positivity Ratio”. In: *Positive psychology*. Ed. by Tim Lomas and Kate Hefferon. Vol. 6. Sage Benchmarks in Psychology. Los Angeles: Sage Reference, pp. 99–125. ISBN: 978-1-4739-0771-3.
- Butler, Julie and Margaret L. Kern (Oct. 13, 2016). “The PERMA-Profiler: A brief multidimensional measure of flourishing”. In: *International Journal of Wellbeing* 6.3, pp. 1–48. ISSN: 11798602. DOI: [10.5502/ijw.v6i3.526](https://doi.org/10.5502/ijw.v6i3.526).

- Cabanas, Edgar and Eva Illouz (2019). *Manufacturing happy citizens: how the science and industry of happiness control our lives*. Medford, MA: Polity Press. ISBN: 978-1-5095-3790-7.
- (2021). *Das Glücksdiktat: und wie es unser Leben beherrscht*. Trans. by Michael Adrian. 2. Auflage. Suhrkamp Taschenbuch 4998. Berlin: Suhrkamp. ISBN: 978-3-518-46998-9.
- Campbell, Angus, Philip E. Converse, and Willard L. Rodgers (1976). *The Quality of American Life: Perceptions, Evaluations, and Satisfaction*. New York: Russell Sage Foundation.
- Campbell, Donald T. and Donald W. Fiske (1959). “Convergent and Discriminant Validation by Multitrait-Multimethod Matrix”. In: *Psychological Bulletin* 56.2, pp. 81–105.
- Chang, Hasok (2004). *Inventing temperature: measurement and scientific progress*. Oxford studies in philosophy of science. Oxford ; New York: Oxford University Press. ISBN: 978-0-19-517127-3.
- (2014). *Is water H<sub>2</sub>O? evidence, realism and pluralism*. Boston studies in the philosophy of science 293. OCLC: ocn870412305. Cambridge, England ; New York: Springer. ISBN: 978-94-007-9646-1.
- (2021). “Operationalism”. In: Zalta, Edward N. *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University.
- Chekola, Mark (1974). *The Concept of Happiness*. Unpublished Dissertation, University of Michigan.
- Costa, Paul T., Robert R. McCrae, and Alan B. Zonderman (Aug. 1987). “Environmental and dispositional influences on well-being: Longitudinal follow-up of an American national sample”. In: *British Journal of Psychology* 78.3, pp. 299–306. ISSN: 00071269. DOI: [10.1111/j.2044-8295.1987.tb02248.x](https://doi.org/10.1111/j.2044-8295.1987.tb02248.x).
- Costa, Paul T. Junior and Robert R. McCrae (1980). “Influence of Extraversion and Neuroticism on Subjective Well-Being: Happy and Unhappy People”. In: *Journal of Personality and Social Psychology* 38.4, pp. 668–678.
- Cronbach, Lee J. and Paul E. Meehl (1955). “Construct Validity in Psychological Tests”. In: *Psychological Bulletin* 52.4, pp. 281–302.
- Csikszentmihalyi, Mihaly (1975). *Beyond boredom and anxiety: Experiencing flow in work and play*. The Jossey-Bass behavioral science series. San Francisco: Jossey-Bass Publishers. ISBN: 978-0-87589-261-0.
- (1990). *Flow: the psychology of optimal experience*. New York: Harper & Row. ISBN: 978-0-06-016253-5.
- Csikszentmihalyi, Mihaly and Jeremy Hunter (2014). “Happiness in Everyday Life: The Uses of Experience Sampling”. In: Csikszentmihalyi, Mihaly. *Flow and the Foundations of Positive Psychology*. Dordrecht: Springer Netherlands, pp. 89–101. ISBN: 978-94-017-9087-1 978-94-017-9088-8. DOI: [10.1007/978-94-017-9088-8\\_6](https://doi.org/10.1007/978-94-017-9088-8_6).

- Csikszentmihalyi, Mihaly and Reed Larson (1987). “Validity and Reliability of the Experience-Sampling Method.” in: *The Journal of Nervous and Mental Disease* 175.9, pp. 526–536. ISSN: 0022-3018. DOI: [10.1097/00005053-198709000-00004](https://doi.org/10.1097/00005053-198709000-00004).
- Danziger, Kurt (1999). “Book Review: Control: A History of Behavioral Psychology”. In: *Canadian Psychology* 40.30, pp. 272–273.
- Davies, William (2015). *The happiness industry: how the government and big business sold us well-being*. London: Verso. ISBN: 978-1-78168-845-8.
- Diener, Ed (1984). “Subjective Well-Being”. In: *Psychological Bulletin* 95.3, pp. 543–575.
- (1994). “Assessing Subjective Well-Being: Progress and Opportunities”. In: *Social Indicators Research* 31, pp. 103–157.
- (1995). “A Value Based Index for Measuring National Quality of Life”. In: *Social Indicators Research* 36, pp. 107–127.
- (2000). “Subjective well-being: The science of happiness and a proposal for a national index.” In: *American Psychologist* 55.1, pp. 34–43. ISSN: 1935-990X, 0003-066X. DOI: [10.1037/0003-066X.55.1.34](https://doi.org/10.1037/0003-066X.55.1.34).
- (Sept. 26, 2006). “One happy autobiography”.
- (Feb. 1, 2007). “Guidelines for National Indicators of Subjective Well-Being and Ill-Being”. In: *Applied Research in Quality of Life* 1.2, pp. 151–157. ISSN: 1871-2584, 1871-2576. DOI: [10.1007/s11482-006-9007-x](https://doi.org/10.1007/s11482-006-9007-x).
- Diener, Ed and Ed Diener (1995). “Cross-Cultural Correlates of Life Satisfaction and Self-Esteem”. In: *Journal of Personality and Social Psychology* 68.4, pp. 653–663.
- Diener, Ed and Robert A. Emmons (1985). “The Independence of Positive and Negative Affect”. In: *Journal of Personality and Social Psychology* 47.5, pp. 1105–1117.
- Diener, Ed, Robert A. Emmons, et al. (Feb. 1985). “The Satisfaction With Life Scale”. In: *Journal of Personality Assessment* 49.1, pp. 71–75. ISSN: 0022-3891, 1532-7752. DOI: [10.1207/s15327752jpa4901\\_13](https://doi.org/10.1207/s15327752jpa4901_13).
- Diener, Ed, Jeff Horwitz, and Robert A. Emmons (1985). “Happiness of the very Wealthy”. In: *Social Indicators Research* 16, pp. 263–274.
- Diener, Ed and Ashgar Iran-Nejad (1986). “The Relationship In Experience Between Various Types of Affect”. In: *Journal of Personality and Social Psychology* 50.5, pp. 1031–1038.
- Diener, Ed and Randy J. Larsen (1993). “The experience of emotional well-being”. In: *Handbook of emotions*. Ed. by M. Lewis and J. M. Haviland. New York: Guilford, pp. 405–415.
- Diener, Ed, Randy J. Larsen, et al. (1985). “Intensity and Frequency: Dimensions Underlying Positive and Negative Affect”. In: *Journal of Personality and Social Psychology* 48.5, pp. 1253–1265.
- Diener, Ed, Ed Sandvik, and William Pavot (1991). “Happiness is the frequency, not the intensity, of positive versus negative affect”. In: *Subjective Well-Being; An Interdisci-*

- plinary Perspective*. Ed. by F. Strack, M. Argyle, and N. Schwarz. New York: Pergamon, pp. 119–136.
- Diener, Ed, Ed Sandvik, William Pavot, and Dennis Gallagher (1991). “Response Artifacts in the Measurement of Subjective Well-Being”. In: *Social Indicators Research* 24.1, pp. 35–56.
- Diener, Ed, Ed Sandvik, Larry Seidlitz, et al. (1993). “The Relationship between Income and Subjective Well-Being: Relative or Absolute?” In: *Social Indicators Research* 28.3, pp. 195–223.
- Diener, Ed, Heidi Smith, and Frank Fujita (1995). “The personality structure of affect”. In: *Journal of Personality and Social Psychology* 69.1, pp. 130–141.
- Diener, Ed, Eunkook Suh, and Shigehiro Oishi (1997). “Recent findings on subjective well-being”. In: *Indian Journal of Clinical Psychology* 24.1, pp. 25–41.
- Diener, Ed, Eunkook M. Suh, Richard E. Lucas, et al. (1999). “Subjective well-being: Three decades of progress”. In: *Psychological Bulletin* 125.2, pp. 276–302.
- Diener, Ed, Eunkook M. Suh, Heidi Smith, et al. (1995). “National Differences in Reported Subjective Well-Being: Why do they occur?” In: *Social Indicators Research* 34, pp. 7–32.
- Easterlin, Richard (1974). “Does Economic Growth Improve the Human Lot? Some Empirical Evidence”. In: *Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz*. Ed. by Paul A. David and Melvin W. Reder. New York: Academic Press, pp. 89–125.
- Ehrenreich, Barbara (2009). *Bright-sided: how the relentless promotion of positive thinking has undermined America*. 1st ed. New York: Metropolitan Books. ISBN: 978-0-8050-8749-9.
- Ehrenreich, Barbara et al. (2010). *Smile or die: wie die Ideologie des positiven Denkens die Welt verdummt*. München: Kunstmann. ISBN: 978-3-88897-682-7.
- Eichstaedt, Johannes C. et al. (Feb. 2015a). “Psychological Language on Twitter Predicts County-Level Heart Disease Mortality”. In: *Psychological Science* 26.2, pp. 159–169. ISSN: 0956-7976, 1467-9280. DOI: [10.1177/0956797614557867](https://doi.org/10.1177/0956797614557867).
- (Feb. 2015b). “Psychological Language on Twitter Predicts County-Level Heart Disease Mortality”. In: *Psychological Science* 26.2, pp. 159–169. ISSN: 0956-7976, 1467-9280. DOI: [10.1177/0956797614557867](https://doi.org/10.1177/0956797614557867).
- Eitler, Pascal and Jens Elberfeld, eds. (2015). *Zeitgeschichte des Selbst: Therapeutisierung, Politisierung, Emotionalisierung*. Histoire Band 79. OCLC: ocn926671512. Bielefeld: Transcript. ISBN: 978-3-8376-3084-8.
- Emmons, Robert A. and Ed Diener (Mar. 1985). “Personality Correlates of Subjective Well-Being”. In: *Personality and Social Psychology Bulletin* 11.1, pp. 89–97. ISSN: 0146-1672, 1552-7433. DOI: [10.1177/0146167285111008](https://doi.org/10.1177/0146167285111008).

- Feest, Uljana (2005). “Operationism in psychology: What the debate is about, what the debate should be about”. In: *Journal of the History of the Behavioral Sciences* 41.2, pp. 131–149. ISSN: 0022-5061, 1520-6696. DOI: [10.1002/jhbs.20079](https://doi.org/10.1002/jhbs.20079).
- (Aug. 30, 2010). “Concepts as Tools in the Experimental Generation of Knowledge in Cognitive Neuropsychology”. In: *Spontaneous Generations: A Journal for the History and Philosophy of Science* 4.1, pp. 173–190. ISSN: 1913-0465. DOI: [10.4245/sponge.v4i1.11938](https://doi.org/10.4245/sponge.v4i1.11938).
- (Dec. 2019). “Why Replication Is Overrated”. In: *Philosophy of Science* 86.5, pp. 895–905. ISSN: 0031-8248, 1539-767X. DOI: [10.1086/705451](https://doi.org/10.1086/705451).
- Feest, Uljana and Friedrich Steinle (2016). “Experiment”. In: *Oxford Handbook of Philosophy of Science*. Ed. by Paul Humphreys. Oxford: Oxford University Press, pp. 274–295.
- Feigl, Herbert (1970). “The “Orthodox” View of Theories: Remarks in Defense as Well as Critique”. In: *Analyses of theories and methods of physics and psychology* 4, pp. 3–16.
- Feldman, Fred (Mar. 18, 2010). *What Is This Thing Called Happiness?* Oxford: Oxford University Press. ISBN: 978-0-19-957117-8.
- Feyerabend, Paul (1982). *Against Method*. 4th ed. London: Verso.
- Fleck, Ludwik (1979). *Genesis and development of a scientific fact*. Ed. by Thaddeus J. Trenn, Robert K. Merton, and Fred Bradley. Sociology of Science. Chicago [u.a]: Univ. of Chicago Press. ISBN: 978-0-226-25325-1.
- (2011). *Denkstile und Tatsachen: gesammelte Schriften und Zeugnisse*. Ed. by Sylwia Werner and Frank Stahnisch. Suhrkamp-Taschenbuch Wissenschaft 1953. Berlin: Suhrkamp. ISBN: 978-3-518-29553-3.
- Fowler, Raymond D., Martin E. P. Seligman, and Gerald P. Koocher (1999). “The APA 1998 Annual Report.” In: *American Psychologist* 54.8, pp. 537–568. ISSN: 0003-066X. DOI: [10.1037/0003-066X.54.8.537](https://doi.org/10.1037/0003-066X.54.8.537).
- Fredrickson, Barbara L. (Sept. 1998). “What Good Are Positive Emotions?” In: *Review of General Psychology* 2.3, pp. 300–319. ISSN: 1089-2680, 1939-1552. DOI: [10.1037/1089-2680.2.3.300](https://doi.org/10.1037/1089-2680.2.3.300).
- (2001). “The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions.” In: *American Psychologist* 56.3, pp. 218–226. ISSN: 1935-990X, 0003-066X. DOI: [10.1037/0003-066X.56.3.218](https://doi.org/10.1037/0003-066X.56.3.218).
- Fredrickson, Barbara L. and Thomas Joiner (Mar. 2002). “Positive Emotions Trigger Upward Spirals Toward Emotional Well-Being”. In: *Psychological Science* 13.2, pp. 172–175. ISSN: 0956-7976, 1467-9280. DOI: [10.1111/1467-9280.00431](https://doi.org/10.1111/1467-9280.00431).
- Fredrickson, Barbara L. and Marcial F. Losada (2005a). “Positive Affect and the Complex Dynamics of Human Flourishing.” In: *American Psychologist* 60.7, pp. 678–686. ISSN: 1935-990X, 0003-066X. DOI: [10.1037/0003-066X.60.7.678](https://doi.org/10.1037/0003-066X.60.7.678).

- Fredrickson, Barbara L. and Marcial F. Losada (2005b). “Positive Affect and the Complex Dynamics of Human Flourishing.” In: *American Psychologist* 60.7, pp. 678–686. ISSN: 1935-990X, 0003-066X. DOI: [10.1037/0003-066X.60.7.678](https://doi.org/10.1037/0003-066X.60.7.678).
- (2013). “”Positive affect and the complex dynamics of human flourishing”: Correction to Fredrickson and Losada (2005).” In: *American Psychologist* 68.9, pp. 822–822. ISSN: 1935-990X, 0003-066X. DOI: [10.1037/a0034435](https://doi.org/10.1037/a0034435).
- Frey, Bruno S. and Alois Stutzer (1999). “Measuring Preferences by Subjective Well-Being”. In: *Journal of Institutional and Theoretical Economics* 155, pp. 755–778.
- (2002a). *Happiness and Economics: How the Economy and Institutions Affect Well-Being*. Princeton, N.J: Princeton University Press. ISBN: 978-0-691-06997-5 978-0-691-06998-2.
- (2002b). *Happiness and Economics: How the Economy and Institutions Affect Well-Being*. Princeton: Princeton University Press. ISBN: 978-0-691-06998-2.
- Froh, Jeffrey J. (2004). “The History of Positive Psychology: Truth Be Told”. In: *NYS Psychologist* 16.3, pp. 18–20.
- Gohm, C. L. et al. (1998). “Culture, parental conflict, parental marital status, and the subjective well-being of young adults”. In: *Journal of Marriage and the Family* 60.2, pp. 319–334.
- Goldings, Herbert Jeremy (Sept. 1954). “On the Avowal and Projection of Happiness”. In: *Journal of Personality* 23.1, pp. 30–47. ISSN: 0022-3506, 1467-6494. DOI: [10.1111/j.1467-6494.1954.tb02336.x](https://doi.org/10.1111/j.1467-6494.1954.tb02336.x).
- Golin, Sanford, Paul D. Sweeney, and David E. Shaeffer (1981). “The causality of causal attributions in depression: A cross-lagged panel correlational analysis.” In: *Journal of Abnormal Psychology* 90.1, pp. 14–22. ISSN: 1939-1846, 0021-843X. DOI: [10.1037/0021-843X.90.1.14](https://doi.org/10.1037/0021-843X.90.1.14).
- Gregersen, Erik and T. Editors of Encyclopaedia (Dec. 23, 2022). *Mohs hardness*. In: *Encyclopædia Britannica*. Encyclopædia Britannica.
- Hartmann, G. W. (July 1934). “Personality traits associated with variations in happiness.” In: *The Journal of Abnormal and Social Psychology* 29.2, pp. 202–212. ISSN: 0096-851X. DOI: [10.1037/h0073108](https://doi.org/10.1037/h0073108).
- Hausman, Daniel M. (2015). *Valuing health: well-being, freedom, and suffering*. Population-level bioethics series. Oxford ; New York: Oxford University Press. ISBN: 978-0-19-023318-1.
- Haybron, Daniel M. (May 2001). “Happiness and Pleasure”. In: *Philosophy and Phenomenological Research* 62.3, pp. 501–528. ISSN: 0031-8205, 1933-1592. DOI: [10.1111/j.1933-1592.2001.tb00072.x](https://doi.org/10.1111/j.1933-1592.2001.tb00072.x).
- (2010). *The pursuit of unhappiness: the elusive psychology of well-being*. OCLC: ocn615830938. Oxford ; New York: Oxford University Press. ISBN: 978-0-19-959246-3.

- Haybron, Daniel M. (2013). *Happiness: a very short introduction*. First edition. Very short introductions 360. OCLC: ocn859188507. Oxford: Oxford University Press. ISBN: 978-0-19-959060-5.
- Haybron, Daniel M. and Valerie Tiberius (2015). “Well-Being Policy: What Standard of Well-Being?” In: *Journal of the American Philosophical Association* 1.4, pp. 712–733. ISSN: 2053-4477, 2053-4485. DOI: [10.1017/apa.2015.23](https://doi.org/10.1017/apa.2015.23).
- Hefferon, Kate and Ilona Boniwell (2011). *Positive psychology: theory, research and applications*. Maidenhead: Open Univ. Press. ISBN: 978-0-335-24194-1 978-0-335-24195-8 978-0-335-24196-5.
- Hoagland, Mahlon B. (1990). *Toward the habit of truth: a life in science*. 1st ed. Commonwealth Fund Book Program. New York: Norton. ISBN: 978-0-393-02754-9.
- Huppert, Felicia A. and Timothy T. C. So (Feb. 2013). “Flourishing Across Europe: Application of a New Conceptual Framework for Defining Well-Being”. In: *Social Indicators Research* 110.3, pp. 837–861. ISSN: 0303-8300, 1573-0921. DOI: [10.1007/s11205-011-9966-7](https://doi.org/10.1007/s11205-011-9966-7).
- Joseph, Jay (2022). “A Reevaluation of the 1990 “Minnesota Study of Twins Reared Apart” IQ Study”. In: *Human Development* 66.1, pp. 48–65. ISSN: 0018-716X, 1423-0054. DOI: [10.1159/000521922](https://doi.org/10.1159/000521922).
- Kleiner, Stephanie (2015). “Staging the Happy Self. Strategies of Epistemic Self-Fashioning by Happiness Researchers (ca. 1950-2000)”. In: *K@iros* 1.
- Kleiner, Stephanie and Robert Suter, eds. (2015). *Guter Rat: Glück und Erfolg in der Ratgeberliteratur 1900-1940*. Glück und Erfolg Band 1. Berlin: Neofelis Verlag. ISBN: 978-3-943414-50-9.
- Kramer, Adam D. I., Jamie E. Guillory, and Jeffrey T. Hancock (June 17, 2014). “Experimental evidence of massive-scale emotional contagion through social networks”. In: *Proceedings of the National Academy of Sciences* 111.24, pp. 8788–8790. ISSN: 0027-8424, 1091-6490. DOI: [10.1073/pnas.1320040111](https://doi.org/10.1073/pnas.1320040111).
- Kringelbach, Morten L. and Kent C. Berridge (2010). “The Neuroscience of Happiness and Pleasure”. In: *Social Research* 77.2, pp. 659–678. ISSN: 0037-783X.
- Kuhn, Thomas S. (1970). *Die Struktur wissenschaftlicher Revolutionen*. Zweite revidierte und um das Postskriptum von 1969 ergänzte Aufl., 24. Auflage. Suhrkamp-Taschenbuch Wissenschaft 25. OCLC: 892436638. Frankfurt am Main: Suhrkamp. ISBN: 978-3-518-27625-9.
- Land, Kenneth C. and Alex C. Michalos (Feb. 2018). “Fifty Years After the Social Indicators Movement: Has the Promise Been Fulfilled?: An Assessment an Agenda for the Future”. In: *Social Indicators Research* 135.3, pp. 835–868. ISSN: 0303-8300, 1573-0921. DOI: [10.1007/s11205-017-1571-y](https://doi.org/10.1007/s11205-017-1571-y).
- Larroulet Philippi, Cristian (Dec. 2021). “On Measurement Scales: Neither Ordinal Nor Interval?” In: *Philosophy of Science* 88.5, pp. 929–939. ISSN: 0031-8248, 1539-767X. DOI: [10.1086/714873](https://doi.org/10.1086/714873).



- Larsen, Randy J. and Ed Diener (1987). "Affect Intensity as an Individual Difference Characteristic: A Review". In: *Journal of Research in Personality* 21, pp. 1–39.
- Larsen, Randy J., Ed Diener, and Robert A. Emmons (1985). "An Evaluation of Subjective Well-Being Measures". In: *Social Indicators Research* 17, pp. 1–17.
- (1986). "Affect Intensity and Reactions to Daily Life Events". In: *Journal of Personality and Social Psychology* 51.4, pp. 803–814.
- Larson, R. (Jan. 1, 1978). "Thirty Years of Research on the Subjective Well-being of Older Americans". In: *Journal of Gerontology* 33.1, pp. 109–125. ISSN: 0022-1422. DOI: [10.1093/geronj/33.1.109](https://doi.org/10.1093/geronj/33.1.109).
- Larson, Reed and Mihaly Csikszentmihalyi (1983). "The Experience Sampling Method". In: *New Directions for Methodology of Social and Behavioral Sciences* 15. Ed. by H.T Reiss, pp. 41–56.
- Latour, Bruno (1986). "Visualisation and Cognition: Drawing Things Together". In: *Knowledge and Society. Studies in the Sociology of Culture Past and Present* 6, pp. 1–40.
- (1987). *Science in Action*. Cambridge: Harvard University Press.
- (1999). "Circulating Reference - Sampling the Soil in the Amazon Forest". In: *Pandora's hope: essays on the reality of science studies*. Cambridge, Mass: Harvard University Press, pp. 25–79. ISBN: 978-0-674-65335-1 978-0-674-65336-8.
- (2012). "The more manipulations the better". In: *New Representations in Scientific Practice*. Ed. by Charlotte Coopmans et al. Cambridge: MIT Press, pp. 158–.
- Layard, Peter R. G. (1980). "Human satisfaction and public policy". In: *Economic Journal* 90.360, pp. 737–750.
- (2005). *Happiness. Lessons from a New Science*. New York: Penguin Books.
- Lazarus, Richard S. (Apr. 2003a). "AUTHOR'S RESPONSE: The Lazarus Manifesto for Positive Psychology and Psychology in General". In: *Psychological Inquiry* 14.2, pp. 173–189. ISSN: 1047-840X, 1532-7965. DOI: [10.1207/S15327965PLI1402\\_04](https://doi.org/10.1207/S15327965PLI1402_04).
- (2003b). "Does the Positive Psychology Movement Have Legs?" In: *Psychological Inquiry* 14, pp. 93–109.
- (2015). "Does the Positive Psychology Movement Have Legs?" In: *Positive psychology*. Ed. by Tim Lomas and Kate Hefferon. Vol. 6. 6 vols. Sage Benchmarks in Psychology. OCLC: ocn881140826. Los Angeles: Sage Reference. ISBN: 978-1-4739-0771-3.
- Leahey, Thomas H. (1992). "Mythical revolutions of American Psychology". In: *American Psychologist* 47.2, pp. 308–318.
- Lee, Yueh-Ting and Martin E.P. Seligman (Jan. 1997). "Are Americans more Optimistic than the Chinese?" In: *Personality and Social Psychology Bulletin* 23.1, pp. 32–40. ISSN: 0146-1672, 1552-7433. DOI: [10.1177/0146167297231004](https://doi.org/10.1177/0146167297231004).
- Leimgruber, Walter (2010). "Überglückt: wohin man blickt, nur Glück". In: *Schweizerisches Archiv für Volkskunde* 106.1, pp. 47–54.
- Lykken, David and Auke Tellegen (1996). "Happiness is a stochastic phenomenon". In: *American Psychological Society* 70.3, pp. 186–189.

- Lyubomirsky, Sonja, Laura King, and Ed Diener (2005). "The Benefits of Frequent Positive Affect: Does Happiness Lead to Success?" In: *Psychological Bulletin* 131.6, pp. 803–855. ISSN: 1939-1455, 0033-2909. DOI: [10.1037/0033-2909.131.6.803](https://doi.org/10.1037/0033-2909.131.6.803).
- Lyubomirsky, Sonja, Kennon M. Sheldon, and David Schkade (June 2005). "Pursuing Happiness: The Architecture of Sustainable Change". In: *Review of General Psychology* 9.2, pp. 111–131. ISSN: 1089-2680, 1939-1552. DOI: [10.1037/1089-2680.9.2.111](https://doi.org/10.1037/1089-2680.9.2.111).
- Lyubomirsky, Sonja, S. Tkach, and Kennon M. Sheldon (2004). "Personality Fit and Positive Interventions: Extraverted and Introverted Individuals Benefit from Different Happiness Increasing Strategies". In: *Unpublished raw data*. DOI: [doi:10.1037/1089-2680.9.2.111](https://doi.org/doi:10.1037/1089-2680.9.2.111).
- Maasen, Sabine et al., eds. (2011). *Das beratene Selbst - Zur Genealogie der Therapeutisierung in den ›langen‹ Siebzigern*. Bielefeld: Transcript Verlag.
- Magnus, Keith et al. (1993). "Extraversion and Neuroticism as Predictors of Objective Life Events: A Longitudinal Analysis". In: *Journal of Personality and Social Psychology* 65.5, pp. 1046–1053.
- Maier, Steven F. and Martin E. P. Seligman (1976). "Learned Helplessness: Theory and Evidence". In: *Journal of Experimental Psychology* 105.1, pp. 3–46.
- Maslow, A. H. (1943). "A theory of human motivation." In: *Psychological Review* 50.4, pp. 370–396. ISSN: 0033-295X. DOI: [10.1037/h0054346](https://doi.org/10.1037/h0054346).
- McAdams, Dan P. et al. (1997). "Stories of commitment: The psychosocial construction of generative lives." In: *Journal of Personality and Social Psychology* 72.3, pp. 678–694. ISSN: 1939-1315, 0022-3514. DOI: [10.1037/0022-3514.72.3.678](https://doi.org/10.1037/0022-3514.72.3.678).
- McMahon, Darrin M. (2006). *Happiness - a History*. New York: Atlantic Monthly Press.
- Michalos, Alex C. (1980). "Review of W. Tatarkiewicz, Analysis of Happiness". In: *Social Indicators Research* 8, pp. 507–510.
- Mills, John A. (1998). *Control: a history of behavioral psychology*. OCLC: 245674527. New York: New York Univ. Press. ISBN: 978-0-8147-5611-9.
- Moneta, Giovanni B. and Mihaly Csikszentmihalyi (June 1996). "The Effect of Perceived Challenges and Skills on the Quality of Subjective Experience". In: *Journal of Personality* 64.2, pp. 275–310. ISSN: 0022-3506, 1467-6494. DOI: [10.1111/j.1467-6494.1996.tb00512.x](https://doi.org/10.1111/j.1467-6494.1996.tb00512.x).
- Myers, D. G. and E. Diener (1995). "Who is happy?" In: *Psychological Science* 6.1, pp. 10–19.
- Nakamura, Jeanne and Mihaly Csikszentmihalyi (2002). "The Concept of Flow". In: *Handbook of positive psychology*. Ed. by C. R. Snyder and Shane J. Lopez. Oxford [England] ; New York: Oxford University Press, pp. 89–105. ISBN: 978-0-19-513533-6.
- Neugarten, B. L., R. J. Havighurst, and S. S. Tobin (Apr. 1, 1961). "The Measurement of Life Satisfaction". In: *Journal of Gerontology* 16.2, pp. 134–143. ISSN: 0022-1422. DOI: [10.1093/geronj/16.2.134](https://doi.org/10.1093/geronj/16.2.134).

- Noll, Heinz-Herbert (2017). “On the History of Social Indicators Research: Measuring, Monitoring and Researching Wellbeing”. In: *Essays by the Members of the Scientific Advisory Board. Government Strategy on Wellbeing in Germany*. Ed. by Federal German Government. Federal Press Office, pp. 6–17.
- Overmier, J. Bruce and Russell C. Leaf (1965). “Effects of discriminative Pavlovian fear conditioning upon previously or subsequently acquired avoidance responding.” In: *Journal of Comparative and Physiological Psychology* 60.2, pp. 213–217. ISSN: 0021-9940.
- Overmier, J. Bruce and Martin E. P. Seligman (1967). “Effects of inescapable shock upon subsequent escape and avoidance responding.” In: *Journal of Comparative and Physiological Psychology* 63.1, pp. 28–33. ISSN: 0021-9940. DOI: [10.1037/h0024166](https://doi.org/10.1037/h0024166).
- Pavot, Bill, Ed Diener, and Frank Fujita (1990). “Extraversion and happiness”. In: *Personality and Individual Differences* 11.12, pp. 1299–1306.
- Peterson, Christopher, Carl von Baeyer, et al. (Sept. 1982). “The attributional Style Questionnaire”. In: *Cognitive Therapy and Research* 6.3, pp. 287–299. ISSN: 0147-5916, 1573-2819. DOI: [10.1007/BF01173577](https://doi.org/10.1007/BF01173577).
- Peterson, Christopher and Martin E. P Seligman (1987). “Explanatory Style and Illness”. In: *Journal of Personality* 55, pp. 237–265.
- (2004). *Character strengths and virtues: a handbook and classification*. Washington, DC : New York: American Psychological Association ; Oxford University Press. ISBN: 978-0-19-516701-6.
- Plutchik, Robert (1980). *Emotion, a psychoevolutionary synthesis: Robert Plutchik*. OCLC: 251297931. New York: Harper & Row. ISBN: 978-0-06-045235-3.
- Popper, Karl R. (2005). *Logik der Forschung*. Ed. by Herbert Keuth. 11. Aufl. durchges. u. erg. Gesammelte Werke in deutscher Sprache / Karl R. Popper. Hrsg.: W. W. Bartley III 3. Tübingen: Mohr Siebeck. ISBN: 978-3-16-148111-6 978-3-16-148410-0.
- (2008). *The Logic of scientific discovery*. Repr. 2008 (twice). Routledge classics. London: Routledge. ISBN: 978-0-415-27843-0 978-0-415-27844-7.
- Rheinberger, Hans-Jörg (1992a). “Experiment, difference, and writing: I. Tracing protein synthesis”. In: *Studies in History and Philosophy of Science Part A* 23.2, pp. 305–331. ISSN: 00393681. DOI: [10.1016/0039-3681\(92\)90037-7](https://doi.org/10.1016/0039-3681(92)90037-7).
- (1992b). “Experiment, difference, and writing: II. The laboratory production of transfer RNA”. In: *Studies in History and Philosophy of Science Part A* 23.3, pp. 389–422. ISSN: 00393681. DOI: [10.1016/0039-3681\(92\)90002-N](https://doi.org/10.1016/0039-3681(92)90002-N).
- (1997). *Toward a history of epistemic things: synthesizing proteins in the test tube*. Writing science. Stanford, Calif: Stanford University Press. ISBN: 978-0-8047-2785-3 978-0-8047-2786-0.
- (2018). “On Science and Philosophy”. In: *Crisis & Critique* 5.1, pp. 341–347.
- Russell, James A and Albert Mehrabian (Sept. 1977). “Evidence for a three-factor theory of emotions”. In: *Journal of Research in Personality* 11.3, pp. 273–294. ISSN: 00926566. DOI: [10.1016/0092-6566\(77\)90037-X](https://doi.org/10.1016/0092-6566(77)90037-X).

- Russell, James A. (1979). "Affective space is bipolar." In: *Journal of Personality and Social Psychology* 37.3, pp. 345–356. ISSN: 0022-3514. DOI: [10.1037/0022-3514.37.3.345](https://doi.org/10.1037/0022-3514.37.3.345).
- Sarason, Irwin G., James H. Johnson, and Judith M. Siegel (1978). "Assessing the impact of life changes: Development of the Life Experiences Survey." In: *Journal of Consulting and Clinical Psychology* 46.5, pp. 932–946. ISSN: 1939-2117, 0022-006X. DOI: [10.1037/0022-006X.46.5.932](https://doi.org/10.1037/0022-006X.46.5.932).
- Schueller, Stephen M. (2012). "Personality Fit and Positive Interventions: Extraverted and Introverted Individuals Benefit from Different Happiness Increasing Strategies". In: *Psychology* 03.12, pp. 1166–1173. ISSN: 2152-7180, 2152-7199. DOI: [10.4236/psych.2012.312A172](https://doi.org/10.4236/psych.2012.312A172).
- Segal, Nancy L. (2012). *Born together-reared apart: The landmark Minnesota twin study*. OCLC: 808085764. Cambridge, Mass: Harvard University Press. ISBN: 978-0-674-05546-9.
- Seidlitz, Larry and Ed Diener (1993). "Memory for Positive Versus Negative Life Events: Theories for the Differences Between Happy and Unhappy Persons". In: *Journal of Personality and Social Psychology* 64.4, pp. 654–664.
- Seligman, Martin E. P. (Feb. 1972). "Learned Helplessness". In: *Annual Review of Medicine* 23.1, pp. 407–412. ISSN: 0066-4219, 1545-326X. DOI: [10.1146/annurev.me.23.020172.002203](https://doi.org/10.1146/annurev.me.23.020172.002203).
- (1975). *Helplessness: on depression, development, and death*. A Series of books in psychology. San Francisco : New York: W. H. Freeman. ISBN: 978-0-7167-0752-3.
- (1991). *Learned optimism*. New York: A.A. Knopf. ISBN: 978-0-394-57915-3.
- (1999a). "Positive Psychology Summit Speech". Lincoln.
- (1999b). "The President's Address (Annual Report)". In: *American Psychologist* 54, pp. 559–562.
- (2002). *Authentic happiness: using the new positive psychology to realize your potential for lasting fulfillment*. New York: Free Press. ISBN: 978-0-7432-2297-6.
- (2011). *Flourish: a visionary new understanding of happiness and well-being*. New York: Free Press. ISBN: 978-1-4391-9075-3.
- Seligman, Martin E. P. and Mihaly Csikszentmihalyi (2000). "Positive psychology: An introduction." In: *American Psychologist* 55.1, pp. 5–14. ISSN: 1935-990X, 0003-066X.
- Seligman, Martin E. P. and Steven F. Maier (1967). "Failure to escape traumatic shock". In: *Journal of Experimental Psychology* 74.1, pp. 1–9.
- Seligman, Martin E. P., Tayyab Rashid, and Acacia C. Parks (Nov. 2006). "Positive psychotherapy." In: *American Psychologist* 61.8, pp. 774–788. ISSN: 1935-990X, 0003-066X. DOI: [10.1037/0003-066X.61.8.774](https://doi.org/10.1037/0003-066X.61.8.774).
- Seligman, Martin E. P., Karen Reivich, et al. (1995). *The optimistic child*. Boston, Mass: Houghton Mifflin. ISBN: 978-0-395-69380-3.
- Seligman, Martin E. P. and Peter Schulman (1986). "Explanatory style as a predictor of productivity and quitting among life insurance sales agents." In: *Journal of Personality*

- and *Social Psychology* 50.4, pp. 832–838. ISSN: 1939-1315, 0022-3514. DOI: [10.1037/0022-3514.50.4.832](https://doi.org/10.1037/0022-3514.50.4.832).
- Sheldon, Ken et al. (2000). “Positive Psychology Manifesto”. In: pp. 1–3.
- Sheldon, Kennon M. and Sonja Lyubomirsky (Mar. 2006). “Achieving Sustainable Gains in Happiness: Change Your Actions, not Your Circumstances\*”. In: *Journal of Happiness Studies* 7.1, pp. 55–86. ISSN: 1389-4978, 1573-7780. DOI: [10.1007/s10902-005-0868-8](https://doi.org/10.1007/s10902-005-0868-8).
- Shin, D. C and D. M. Johnson (1978). “Avowed Happiness as an Overall Assessment of the Quality of Life”. In: *Social Indicators Research* 5, pp. 475–492.
- Sirgy, M. Joseph et al. (May 2006). “The Quality-of-Life (QOL) Research Movement: Past, Present, and Future”. In: *Social Indicators Research* 76.3, pp. 343–466. ISSN: 0303-8300, 1573-0921. DOI: [10.1007/s11205-005-2877-8](https://doi.org/10.1007/s11205-005-2877-8).
- Skinner, B. F. (1957). *Verbal Behavior*. Ed. by Richard M. Elliott and Kenneth MacCorquodale. The Century Psychology Series. New York: Appleton-Century-Crofts. ISBN: 57-13446.
- (1976). *About behaviorism*. New York: Vintage Books. ISBN: 978-0-394-71618-3.
- Smith, Laurence D. (1986). *Behaviorism and logical positivism: a reassessment of the alliance*. Stanford, Calif: Stanford University Press. ISBN: 978-0-8047-1301-6.
- Spielman, Rose M. et al. (2017). *Psychology*. OCLC: 900470589. Houston, Texas: OpenStax College, Rice University. ISBN: 978-1-947172-07-4.
- Staw, Barry M., Robert I. Sutton, and Lisa H. Pelled (Feb. 1994). “Employee Positive Emotion and Favorable Outcomes at the Workplace”. In: *Organization Science* 5.1, pp. 51–71. ISSN: 1047-7039, 1526-5455. DOI: [10.1287/orsc.5.1.51](https://doi.org/10.1287/orsc.5.1.51).
- Stevens, S. S. (1946). “On the Theory of Scales of Measurement”. In: *Science* 103.2684, pp. 677–680.
- Suppes, Patrick and Joseph L. Zinnes (1963). “Basic measurement theory”. In: *Handbook of Mathematical Psychology*. Ed. by D. Luce and Robert Bush. Vol. 36. London: John Wiley & Sons, Inc, pp. 1–76.
- Symonds, P. M. (1937). “Happiness as related to problems and interests.” In: *Journal of Educational Psychology* 28.4, pp. 290–294. ISSN: 0022-0663. DOI: [10.1037/h0061067](https://doi.org/10.1037/h0061067).
- Tal, Eran (Dec. 2013). “Old and New Problems in Philosophy of Measurement: Old and New Problems in Philosophy of Measurement”. In: *Philosophy Compass* 8.12, pp. 1159–1173. ISSN: 17479991. DOI: [10.1111/phc3.12089](https://doi.org/10.1111/phc3.12089).
- (2020). *Measurement in Science*. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Zalta N. Edward. Fall 2020. Stanford: Metaphysics Research Lab.
- Taylor, Eugene (Jan. 1, 2001). “Positive Psychology and Humanistic Psychology: A Reply to Seligman”. In: *Journal of Humanistic Psychology* 41.1, pp. 13–29. ISSN: 0022-1678. DOI: [10.1177/0022167801411003](https://doi.org/10.1177/0022167801411003).
- Thomas, David L. and Ed Diener (1990). “Memory accuracy in the recall of emotions”. In: *Journal of Personality and Social Psychology* 59.2, pp. 291–297.

- Tiberius, Valerie (2006). “Well-Being: Psychological Research for Philosophers”. In: *Philosophy Compass* 1.5, pp. 493–505. DOI: <https://doi.org/10.1111/j.1747-9991.2006.00038.x>.
- (2010). *The reflective life: living wisely with our limits*. 1. publ. in paperback. Oxford: Oxford University Press. ISBN: 978-0-19-957560-2.
- (2015). “Well-being, Values and Improving Lives”. In: *Performance and Progress: Essays on Capitalism, Business, and Society*. Ed. by Subramanian Rangan. Oxford University Press. ISBN: 978-0-19-874428-3. DOI: [10.1093/acprof:oso/9780198744283.001.0001](https://doi.org/10.1093/acprof:oso/9780198744283.001.0001).
- (2018). *Well-being as value fulfillment: how we can help each other to live well*. First edition. OCLC: on1028600955. Oxford University Press: Oxford University Press. ISBN: 978-0-19-880949-4.
- Tiberius, Valerie and Alexandra Plakias (2010). “Well-Being”. In: *The Moral Psychology Handbook*. Ed. by John M. Doris and Moral Psychology Research Group. Oxford University Press. ISBN: 978-0-19-958214-3. DOI: [10.1093/acprof:oso/9780199582143.001.0001](https://doi.org/10.1093/acprof:oso/9780199582143.001.0001).
- U.S. Department of Health Education, {and} Welfare (1969). *Toward a social report*. Washington, D.C.: U.S. Government Printing Office.
- Van Fraassen, Bas C. (2008). *Scientific representation: paradoxes of perspective*. OCLC: ocn216938527. Oxford : New York: Clarendon Press ; Oxford University Press. ISBN: 978-0-19-927822-0.
- Veenhoven, R (1984). *Conditions of happiness*. OCLC: 10727064. Dordrecht, Holland; Boston; Hingham, MA, U.S.A.: D. Reidel. ISBN: 978-90-277-1792-4 978-90-277-1794-8.
- Veenhoven, Ruut (2014). “Journal of Happiness Studies”. In: *Encyclopedia of Quality of Life and Well-Being Research*. Ed. by Alex C. Michalos. Dordrecht: Springer Netherlands, pp. 3462–3464. ISBN: 978-94-007-0752-8 978-94-007-0753-5. DOI: [10.1007/978-94-007-0753-5\\_4049](https://doi.org/10.1007/978-94-007-0753-5_4049).
- Vessonen, Elina (Oct. 2019). “Operationalism and realism in psychometrics”. In: *Philosophy Compass* 14.10. ISSN: 1747-9991, 1747-9991. DOI: [10.1111/phc3.12624](https://doi.org/10.1111/phc3.12624).
- (July 14, 2021a). “Conceptual engineering and operationalism in psychology”. In: *Synthese*. ISSN: 0039-7857, 1573-0964. DOI: [10.1007/s11229-021-03261-x](https://doi.org/10.1007/s11229-021-03261-x).
- (Sept. 2021b). “Representation in measurement”. In: *European Journal for Philosophy of Science* 11.3, p. 76. ISSN: 1879-4912, 1879-4920. DOI: [10.1007/s13194-021-00365-6](https://doi.org/10.1007/s13194-021-00365-6).
- Watson, David, Lee Anna Clark, and Auke Tellegen (1988). “Development and validation of brief measures of positive and negative affect: The PANAS scales.” In: *Journal of Personality and Social Psychology* 54.6, pp. 1063–1070. ISSN: 1939-1315, 0022-3514. DOI: [10.1037/0022-3514.54.6.1063](https://doi.org/10.1037/0022-3514.54.6.1063).
- Watson, John B. (1913). “Psychology as the behaviorist views it.” In: *Psychological Review* 20.2, pp. 158–177. ISSN: 1939-1471, 0033-295X. DOI: [10.1037/h0074428](https://doi.org/10.1037/h0074428).

Wilson, W. R. (1960). *An attempt to determine some correlates and dimensions of hedonic tone*. Unpublished doctoral dissertation, Northwestern University.

Wilson, Warner R. (1967). "Correlates of avowed happiness." In: *Psychological Bulletin* 67.4, pp. 294-306. ISSN: 1939-1455, 0033-2909. DOI: [10.1037/h0024431](https://doi.org/10.1037/h0024431).

## List of Figures

1	A section of a Twitter wordcloud, correlating emotional word use on Twitter with data on heart disease (Eichstaedt et al. 2015b, p. 158). <sup>88</sup> . . . . .	6
2	Listing of nations taking part in the European Social Survey according to the operational definition of "flourishing" (Huppert and So 2013, p. 848). . . . .	12
1.1	Introductory illustration from "The More Manipulations the Better" (Latour 2012, p. 158). . . . .	24
1.2	Stevens scales (Stevens 1946, p. 678). . . . .	29
2.1	Questionnaire item measuring happiness according to (Bradburn and Caplovitz 1965, p. 178). <sup>89</sup> . . . . .	38
2.2	Bradburn's items measuring affect (Bradburn and Caplovitz 1965, p. 17), yellow emphasis by me. . . . .	38
2.3	The seven-point D-T Scale (Andrews and Withey 1976, p. 18). . . . .	41
2.4	The core cluster (encircled) of the life-as-a-whole responses. Each circle with a "G-" title symbolizes the results of a single- or multi-item questionnaire (Andrews and Withey 1976, p. 82). . . . .	43
2.5	The model of subjective-well-being before Diener (P. T. J. Costa and McCrae 1980, p. 675). . . . .	45
3.1	"A Rudimentary Measurement Model" by Andrews and Withey (1976, p. 181). . . . .	54
3.2	Factor Analyses of Adjectives Within the Discrete Emotion Categories (Ed Diener, H. Smith, and Fujita 1995, p. 134). . . . .	62
3.3	Multitrait-Multimethod Matrix of PA (MTMM) (Ed Diener, H. Smith, and Fujita 1995, p. 134). . . . .	63
3.4	Two-factor model of affect structure by (Ed Diener, H. Smith, and Fujita 1995, p. 137). . . . .	63
4.1	The "shuttle box" within which the dog experiments took place (see Overmier and Leaf 1965; Martin E. P. Seligman and Maier 1967), illustration from (Spielman et al. 2017, p. 519). . . . .	73
4.2	The three primary factors that determine "chronic" happiness levels (Lyubomirsky, K. M. Sheldon, and Schkade 2005, p. 116). . . . .	77



5.1	Representation of the dynamic structures describing high-performance, medium-performance, and low-performance teams (Fredrickson and Losada 2005b, p. 682). . . . .	85
5.2	Feigl's representation the connection between postulates and observations (Feigl 1970, p. 6). . . . .	87