Temporal and Population Dynamics of Depressive Symptoms: Empirical and Modeling Approaches

Tom Rosenström

Institute of Behavioural Sciences,
University of Helsinki, Finland

Academic Dissertation to be publicly discussed,
by due permission of the Faculty of Behavioural Sciences,
at the University of Helsinki, Main Building, Auditorium XV, Unioninkatu 34,
on the 19th October, at 10 o'clock

University of Helsinki
Institute of Behavioural Sciences
Studies in Psychology, 95, 2013

Supervisors:

Professor Liisa Keltikangas-Järvinen
Institute of Behavioural Sciences, University of Helsinki, Finland

Docent Mirka Hintsanen
Institute of Behavioural Sciences, University of Helsinki, Finland

Assistant Professor Markus Jokela
Institute of Behavioural Sciences, University of Helsinki, Finland

Reviewers:

Professor Peter Molenaar
Department of Human Development and Family Studies, The Pennsylvania State University, United States of America

Docent Sami Pirkola
Department of Psychiatry, Helsinki University Central Hospital, Finland

Opponent:

Associate Professor Edward H. Hagen
Department of Anthropology, Washington State University, Vancouver, United States of America

ISSN-L 1798-842X

ISSN 1798-842X

ISBN 978-952-10-9338-8 (pbk.)

ISBN 978-952-10-9339-5 (PDF)

http://ethesis.helsinki.fi

Helsinki University Print

Helsinki 2013

CONTENTS

ABSTRACT

Depression has been estimated to be the second largest cause of years lived with a disability, and much research on depressive symptoms exists. Despite this, basic research has not found natural taxa that would correspond to clinical diagnoses for depression. It is often assumed that a one-dimensional latent continuum underlies depressive symptoms, but empirical evidence does not support this idea either. Therefore, it has been suggested that depressive symptoms are part of a complex causal network that has not yet been adequately understood.

This thesis aims to understand statistical variation and joint variation of individual depressive symptoms over time, the causal relationships between these symptoms, and their potentially adaptive evolutionary origins. The research material consists of the prospective Young Finns study that began in 1980 and included 3596 participants, the 10 317-participant Wisconsin Longitudinal study, and mathematical models for the evolution of cooperation.

First, temporal trajectories of a depressive-symptom sum was modeled with a stochastic differential equation model which results in a more empirically justified approach than typical latent-trait models, allowing causal roles for individual symptoms rather than viewing them as passive reflections. Regarding individual symptoms, it was shown that body-image dissatisfaction was the most temporally stable symptom, and strongly associated with chronically elevated dysphoria over a 16 year follow-up. In contrast, symptoms related to sleep and tiredness were the least stable, and novel methods based on non-Gaussian distributions suggested that sleep problems cause other depressive symptoms. Finally, combining the bargaining models of depression with mathematical models for the evolution of cooperation showed that, in theory, evolution should favor the emergence of depressive symptoms in natural populations, as they promote fitness-enhancing cooperation by rendering defection from joint enterprises less tempting. Overall, instead of a single disorder, depressive symptoms may reflect multiple processes, some of them being adaptive instead of dysfunctional.

TIIVISTELMÄ

Masennus on arvioitu toisiksi suurimmaksi toimintakyvyttömänä elettyjen elinvuosien aiheuttajaksi maailmassa, ja masennusoireita on tutkittu paljon. Tästä huolimatta perustutkimus ei ole löytänyt kliinistä diagnostiikkaa vastaavaa luonnollista rajaa masentuneiden ja ei-masentuneiden välille. Yksiulotteinen jatkumo usein oletetaan masennusoireiden taustavaikuttajaksi, mutta tätäkään ajatusta uusin tutkimus ei tue. Oireiden ajatellaan olevan osa huonosti tunnettua syy- ja seuraussuhteiden verkostoa.

Tässä väitöskirjassa pyritään ymmärtämään yksittäisten masennusoireiden tilastollista vaihtelua ja yhteisvaihtelua ajassa, oireiden kausaaliyhteyksiä, ja niiden mahdollisia evoluution kannalta adaptiivisia ominaisuuksia. Tutkimusmateriaalina toimivat 1980-luvulta asti seurattu 3596:n suomalaisen havaintoaineisto, Lasten Sepelvaltimotaudin Riskitekijät -tutkimus, 10 317 amerikkalaisen havaintoaineisto, Wisconsinin pitkättäistutkimus, sekä yhteistyön evoluution matemaattiset mallit.

Usein tutkittua masennusoiresummaa tarkasteltiin stokastisen differentiaaliyhtälömallin avulla. Se tuottaa tyypillisiä latentin taustatekijän malleja realistisemman tilastollisen aikasarjakuvauksen, ja mahdollistaa yksittäisten oireiden toimimisen kausaalisesti aktiivisina tekijöinä yksiulotteisen taustatekijän heijastusten sijaan. Näistä yksittäisistä oireista todettiin, että tyytymättömyys omaan ruumiinkuvaan oli sekä ajassa poikkeuksellisen pysyvää että vahvasti yhteydessä 16 vuoden ajan suhteellisen korkeana pysyneeseen masennuspisteeseen. Sen sijaan uneen ja väsymykseen liittyvä oireilu oli lyhytkestoisempaa, ja uudenlaisen epä-Gaussisiin jakaumiin perustuvan päättelyn mukaan aiheutti muita masennusoireita. Yhdistämällä masennuksen neuvottelustrategia-malli yhteistyön evoluution matemaattisiin malleihin osoitettiin että teoriassa evoluution tulisi suosia masennuspiirteiden kehittymistä populaatiotasolla. Ne voivat edesauttaa yhteistyön syntymistä vähentämällä vapaamatkustamisen yksilökohtaisia hyötyjä. Tulokset vihjaavat että yksittäisen häiriön sijaan, masennusoireet heijastelevat useita prosesseja, joista osa voi olla ennemmin adaptiivisia kuin toimintahäiriöitä.

ACKNOWLEDGEMENTS

First and foremost, I thank my supervisor, Professor Liisa Keltikangas-Järvinen, for making this thesis possible. She has created an incredibly multidisciplinary research group that promotes scientific diversity. She always encourages novel ideas, and her positive support helped greatly when some of the ideas did not take off. I am also grateful to my other supervisors, Assistant Professor Markus Jokela and Docent Mirka Hintsanen, who provided important contributions throughout the research process, and guided me into the art of scientific publishing.

I also thank the Young Finns study researchers, whose ongoing and persistent work made this thesis a possibility, and the reviewers of my dissertation, Professor Peter Molenaar and Docent Sami Pirkola, who provided valuable comments at the very end of the thesis project. Throughout the process, I was supported by stimulating conversations with colleagues, many whom are also co-authors: Doctors Marius Lahti and Kim Josefsson, Christian Hakulinen, Ilmari Määttänen, Docents Laura Pulkki-Råback, Sampsa Puttonen, and Taina Hintsa, Doctors Päivi Merjonen and Sari Mullola, Professors Marko Elovainio, Mika Kivimäki, and C. Robert Cloninger, Aki Kuitunen, and many others. Thank you. Special thanks to Miika Pihlaja and Lauri Nurminen for sharing the rewarding but intimidating initial journeys to mathematics, to Docent Dario Gasbarra for the guidance and the wealth of information (the books!), and to Docent Simo Vanni for the preparing experiences in rigorous imaging neuroscience.

This study was conducted at the Institute of Behavioural Sciences, University of Helsinki, and I thank the Research Foundation of the University of Helsinki, the Oskar Öflund Foundation, and the Alfred Kordelin Foundation for directly funding my research, and the Academy of Finland for indirect funding, and the National Doctoral Programme of Psychology for granting me an affiliate position and a travel grant.

I am always indebted by the love and support of my parents Anja and Leif, brother Jan, and that of the love of my life, Betty. Thank you for being there and for helping me to pursue my goals. Also, thanks to my friends from the very first classes.

LIST OF ORIGINAL CONTRIBUTIONS

I. Rosenström, T., Jokela, M., Hintsanen, M., Pulkki-Råback, L., Hutri-Kähönen, N., Keltikangas-Järvinen, L. (2013). Longitudinal course of depressive symptoms in adulthood: stochastic differential equation modeling. *Psychological Medicine, 43*, 933–944.

II. Rosenström, T., Jokela, M., Hintsanen, M., Josefsson, K., Juonala, M., Kivimäki, M., Pulkki-Råback, L., Viikari, J. S., Hutri-Kähönen, N., Heinonen, E., Raitakari, O. T., Keltikangas-Järvinen, L. (2013). Body-image dissatisfaction is strongly associated with chronic dysphoria. *Journal of Affective Disorders, 150*, 253–260.

III. Rosenström, T., Jokela, M., Puttonen, S., Hintsanen, M., Pulkki-Råback, L., Viikari, J. S., Raitakari, O. T., Keltikangas-Järvinen, L. (2012). Pairwise measures of causal direction in the epidemiology of sleep problems and depression. *PLoS One, 7*, e50841.

IV. Rosenström, T. (2013). Bargaining models of depression and evolution of cooperation. *Journal of Theoretical Biology, 331*, 54–65.

LIST OF ABBREVIATIONS

AIC	Akaike's Information Criterion
BDI	Beck's Depression Inventory
BIC	Bayesian Information Criterion
CES-D	Center for Epidemiologic Studies Depression Scale
DSM	Diagnostic and Statistical Manual of Mental Disorders
DynAffect	Dynamics of Affect Model
EDM	Exact Discrete Model
FIML	Full Information Maximum Likelihood
ICD	International Classification of Diseases
LiNGAM	Linear Non-Gaussian Acyclic Model
mBDI	Modified Beck's Depression Inventory
MDD	Major Depressive Disorder
SES	Socioeconomic Status

1 INTRODUCTION

Depressive disorders have been around throughout the recorded history: “there are few psychiatric syndromes whose clinical descriptions are so constant through successive eras of history” (Beck, 1967). Today, Major Depressive Disorder (MDD) is the second greatest contributor to global years lived with a disability (Vos et al., 2013). It is fair to say that depression has become one of the great public-health puzzles of our time. Depression is defined by its symptoms, and the broad aim of this thesis is to investigate the dynamics and evolution of depressive symptoms in general human populations; but first, some more numbers regarding the public-health significance of those symptoms are provided.

Among categories of diseases and disorders, “mental and behavioral disorders” are the largest current global contributor to years of life lived with a disability (Vos et al., 2013). Among the mental and behavioral diseases, MDD is the single largest contributor to global disability burden; and among all diseases and disorders, MDD is second only to Low back pain (Vos et al., 2013). Depressed persons have much lower subjective health status compared to the non-depressed general population with natural prevalence of other diseases and disorders (Bromet et al., 2011). The magnitude of this effect is 1.4 standard deviations in the World Health Organization’s normative disability scale (Bromet et al., 2011), when already differences of 0.8 standard deviations are generally viewed as large-magnitude differences (Cohen, 1988). In addition to the years lived with a disability, depressive disorders are associated with increased mortality, medical expanses, work absenteeism, and reduction of productivity (Benden & Farvolden, 2010; Karpansalo et al., 2005; Mykletun et al., 2009).

Depression has been estimated to imply an increase in mortality risk as great as smoking, even when adjusting for concurrent smoking (Mykletun et al., 2009). The increased mortality among the depressed people compared to other population is not

simply explained by somatic problems (Mykletun et al., 2009). Neither does the symptom of suicidality explain increased mortality among depressed people (Davies, Naik, & Lee, 2001). Depression increases risk of committed suicide, but it also increases mortality rate for all major disease-related causes of death, and its potential effect on these other causes accounts for a far higher number of deaths than suicides do (Mykletun et al., 2007). Mechanisms underlying the depression-mortality association are not well understood, and complex bi-directional relationships between depression and physical diseases may exist (de Jonge & Roest, 2012; Mykletun et al., 2009).

The estimated global prevalence of MDD is 4.33% of the world's population, being clearly higher in women (5.48%) than in men (3.21%) (Vos et al., 2013). Numbers are similar in Finland, where pure depressive disorder has been found from 4.4% of the population aged 30 or more, and comorbid depression from 6.5% (Pirkola et al., 2005). Large country-wise differences in prevalence estimates exist, however, ranging from 2.2% in Japan to 10.4% in São Paulo, Brazil; although these cross-national comparisons are subject to several methodological problems, and probably biased towards too conservative estimation (Bromet et al., 2011). Depression also runs in the family to some extent, with heritability estimates ranging from 31% to 42% of the total between-individual variance (Bienvenu, Davydow, & Kendler, 2011; Sullivan, Neale, & Kendler, 2000). A major factor that hinders the understanding of depressive disorders is the difficulty of accurately defining them (Ghaemi, Vöhringer, & Whitham, 2013; van Loo, de Jonge, Romeijn, Kessler, & Schoevers, 2012).

1.1 Depressive symptoms

Medical care and associated reimbursement systems have been built around the concept of diagnosis. One either has or does not have a certain disease or disorder. This approach results in binary (yes/no) definitions of depression. Currently, there are two widely used systems for binary diagnostic definitions of depression, one based on the Diagnostic and Statistical Manual of Mental Disorders IV (DSM-IV), and another based

on International Classification of Diseases 10 (ICD-10). DSM is published by the American Psychiatric Association (American Psychiatric Association, 2000), whereas ICD is the World Health Organization's standard diagnostic tool for epidemiology, health management, and clinical purposes (World Health Organization, 2010).

ICD-10 defines that a mild *depressive episode* can be diagnosed if the patient is distressed and has *two or three* of the following symptoms: lowering of mood that varies little from day to day and is unresponsive to circumstances; reduction of energy; decrease in activity; reduced capacity for enjoyment; reduced capacity for interest; reduced capacity for concentration; marked tiredness after minimum effort; disturbed sleep; waking in the morning several hours before the usual time, with depression worst in morning; diminished appetite; reduced self-esteem and -confidence; ideas of guilt or worthlessness; psychomotor retardation; psychomotor agitation; weight loss; and loss of libido. A moderate episode is implicated by four or more symptoms and great difficulty in continuing with daily activities. A severe episode implies further marked distress, and suicidal thoughts and somatic symptoms are often present. Exclusion criteria are simultaneous diagnosis for adjustment disorder, recurrent depressive disorder (see below), or conduct disorder.

ICD-10 also recognizes *recurrent depressive disorder*, which is a disorder characterized by repeated episodes of depression as described by the *depressive episodes* diagnoses (there are several alike to the above-described), without history of independent episodes of mood elevation and overactivity, as in mania. Before recovering from recurrent depressive disorder, a person can be diagnosed as being in *remission* phase, which means that the person has had two or more episodes in the past, but has been free of the symptoms for several months. Returning of the depressive-episode symptoms is called a *relapse*.

DSM-IV is otherwise quite similar to ICD-10, but states that at least one of the two core symptoms (depressed mood and/or loss of interest in virtually all activities) must have been present a minimum of two weeks to warrant a diagnosis, and altogether at least five symptoms must be present for two weeks to warrant a diagnosis of major depressive episode and associated *Major Depressive Disorder* (MDD). Exclusion

criteria can preclude MDD diagnosis: symptoms must cause significant distress or impairment in functioning, they ought not be caused by substance use or a general medical condition, nor occur due to bereavement associated with a recently lost bond (Zimmerman, McGlinchey, Chelminski, & Young, 2006). DSM-IV recognizes nine depressive symptoms, and therefore these probably are the most studied ones globally (Table 1 in the section 1.4 summarizes the symptoms).

Although DSM and ICD systems have facilitated data collection, basic-research studies have not found solid evidence for discrete clustering of depressive symptoms (Haslam, Holland, & Kuppens, 2012; Jokela et al., 2011; Solomon, Haaga, & Arnow, 2001; van Loo et al., 2012). Regarding general measures of psychopathology, use of continuous measures has been estimated to provide an expected 15% increase in reliability and 37% increase in validity compared to discrete measures (Markon, Chmielewski, & Miller, 2011). Depressive symptoms below the five symptoms required for diagnosis of MDD associate with significant functional impairment, the impairment being much more tightly related to degree of depressive-symptoms severity than to diagnostic symptom counts (Karsten, Hartman, Ormel, Nolen, & Penninx, 2010). For the time being, depressive symptoms appear as a continuously varying phenomena in population—a matter of degree rather than a clear-cut, well-specified, condition.

A typical 'continuous' depressive-symptoms score is formed by summing numeric questionnaire answers for several items. Each item yields a progressively higher score based on the severity, frequency, or description-to-person fit that the respondent assigns for the given item. Such sum scores can be interpreted either as convenient summaries of symptom burden or as reflecting a latent causal factor. Latent factors are deeply ingrained in the psychometric literature (Borsboom, Mellenbergh, & Heerden, 2003; Lawley & Maxwell, 1971; Lord & Novick, 1968; Molenaar, 2004); they have been applied a lot in depression research, but have recently fallen under critique (Cramer, Borsboom, Aggen, & Kendler, 2012; Cramer, Waldorp, van der Maas, & Borsboom, 2010). Depressive symptoms cannot be readily decomposed into a single common (latent) factor plus unique variances independent to individual symptoms. So far researchers have been able to reduce the common variance of depressive symptoms only

into 2 to 7 factors, with average over studies being 3.5 factors (Shafer, 2006; van Loo et al., 2012), and the obtained factors have involved inconsistent symptom patterns across the studies (van Loo et al., 2012). Even though we could observe that between-individual covariance in depressive symptoms falls onto single dimension (which we cannot), this would not logically imply homogeneity of the within-individual processes leading to depression, nor existence of a single causal entity behind depression (Borsboom et al., 2003; Molenaar & Campbell, 2009; Molenaar, 2004; van der Maas et al., 2006). Indeed, the time-course (Jokela et al., 2011) and inter-correlation patterns (Cramer et al., 2012) of depressive symptoms appear to vary as a function of life events, and multiple other biological and cognitive factors may affect their development (Hyde, Mezulis, & Abramson, 2008; Kendler, Gardner, & Prescott, 2006; Kendler & Gardner, 2010).

In an attempt to explain the lack of fit for traditional psychometric models, it has been suggested that the depressive symptoms form a causal network where activating one symptom can promote the future emergence of the other symptoms (Cramer et al., 2012). This hypothesis is consistent, for example, with the observations that sleep restriction induces also other depressive symptoms than sleep problems in healthy non-depressed persons, and that insomnia is a risk factor for future depression (Almeida, Alfonso, Yeap, Hankey, & Flicker, 2011; Paunio et al., 2009; Wiebe, Cassof, & Gruber, 2012). Such findings imply that sleep problems do not simply reflect a latent cause for depression, but are a causally active component themselves.

The starting point of this thesis is that the linear latent-factor interpretation (reflective model) of depression is too simplistic, but also that studying of sum-scores is nonetheless useful due to the accumulated empirical evidence regarding the highly inter-correlated depressive symptoms. A wealth of information now exists on the variance in sum scores and diagnoses shared by twins (Kendler et al., 2011), by genome-wide single-nucleotide polymorphisms in population (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013a), by individual polymorphisms, and by polymorphisms related to other major neuropsychiatric disorders (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013a, 2013b), and on the

environmental etiology (*e.g.*, Kendler & Gardner, 2010, 2011). The aim is to embrace the empirically observed complexity in depression research (*e.g.*, Cramer et al., 2012; Kendler & Gardner, 2010, 2011) instead of celebrating individual breakthroughs, however.

Four relatively novel approaches for the study of depressive symptoms are provided, and motivated by previous empirical observations. In order to understand the natural course of depressive symptoms, this thesis studies empirical general-population samples that provide inferences from the naturally occurring statistical distributions rather than clinical populations whose distributions are subject to complex selection processes (availability of treatment, treatment tradition, local specialization, comorbidity, and so on).

1.2 Time trajectories of sum scores

Although twin studies have attributed 58%-67% of the population variance in depression to individual-specific environmental effects and/or measurement error (Sullivan et al., 2000), genetic research has dealt mainly with the presence of depression and not with its course over time. Several different environmental exposures can alter the course of depressive symptoms (Cramer et al., 2012; Jokela et al., 2011; Kendler, Karkowski, & Prescott, 1998). The extent of this dynamic nature of the time course of depressive-symptoms scores in non-clinical population samples remain poorly understood, however. The first aim of this thesis was therefore to estimate the extent to which the variation in sum of depressive symptoms in adulthood reflects (1) temporally stable, trait-like, differences between individuals, (2) the effects of protective and risk factors accumulating over time, and (3) random and non-systematic state fluctuations or measurement error. Such a variance decomposition would allow one, for example, to assess the hypothesis that women encounter more factors than men that together ‘push and pull’ the depressive symptoms over time; women are known to have more variance in, and higher average levels of, depressive symptoms (Hyde et al., 2008).

At least one study has estimated a similar model for children and adolescents followed for three years, with a half-year intervals. Using a structural equation model (Bollen, 1989; Kenny & Zautra, 1995), children and adolescent depression was decomposed into three components similar to those described above [*i.e.*, trait, accumulating influences, and state variation (Cole & Martin, 2005)]. The second component, time-accumulation, was described with an autoregressive process; it is distinct from the static decompositions usually seen in genetic and measurement-related studies, as it adds an element of time evolution to the model. Importantly, an autoregressive process can contain randomness, and thereby model different trajectories for different observed individuals instead of just group-level time evolution. Unfortunately, autoregressive processes are problematic for longitudinal epidemiologic data sets, as they are not readily interpretable across the heterogeneous measurement-time intervals commonly encountered [see first panel of Figure 1 and prior research (Bartlett, 1946; Oud & Delsing, 2010; Oud & Jansen, 2000)]. Fortunately, there is a continuous-time process that can model unobserved random factors accumulating in time.

Brownian motion was named after botanist Robert Brown (1773-1858), who observed water-suspended particles from pollen grain through the microscope, noticing that the particles moved through the water. Brown was unable to determine the mechanism of the motion, but Albert Einstein (1879-1955) showed in the year 1905 that it could be explained by collisions with even smaller unobserved particles, thus providing some of the first historical evidence for the existence of atoms (Einstein, 1956). Brownian motion is also known as the Wiener process, in honor of Norbert Wiener (1894-1964) who established the solid mathematical basis; strictly speaking, the Wiener process is the mathematical model for the phenomenon of Brownian motion. The Wiener process is among the simplest continuous-time stochastic processes, and an end-result of many limits of sequences of both simpler and more complex processes [see random-walk and Donsker's theorem (Klenke, 2008)]; it is the most often applied random element in the theory of stochastic (*i.e.*, randomness-involving) differential equations (Klenke, 2008; Øksendal, 2003). The Wiener process is a model for a time

trajectory of an agent or object that is continuously perturbed by many small influences (second panel of Figure 1 shows two example paths of one-dimensional Brownian motion). Therefore, it is a viable choice for modeling the 'push and pull' that many unobserved factors may cause in depression sum scores of the epidemiologic data sets.

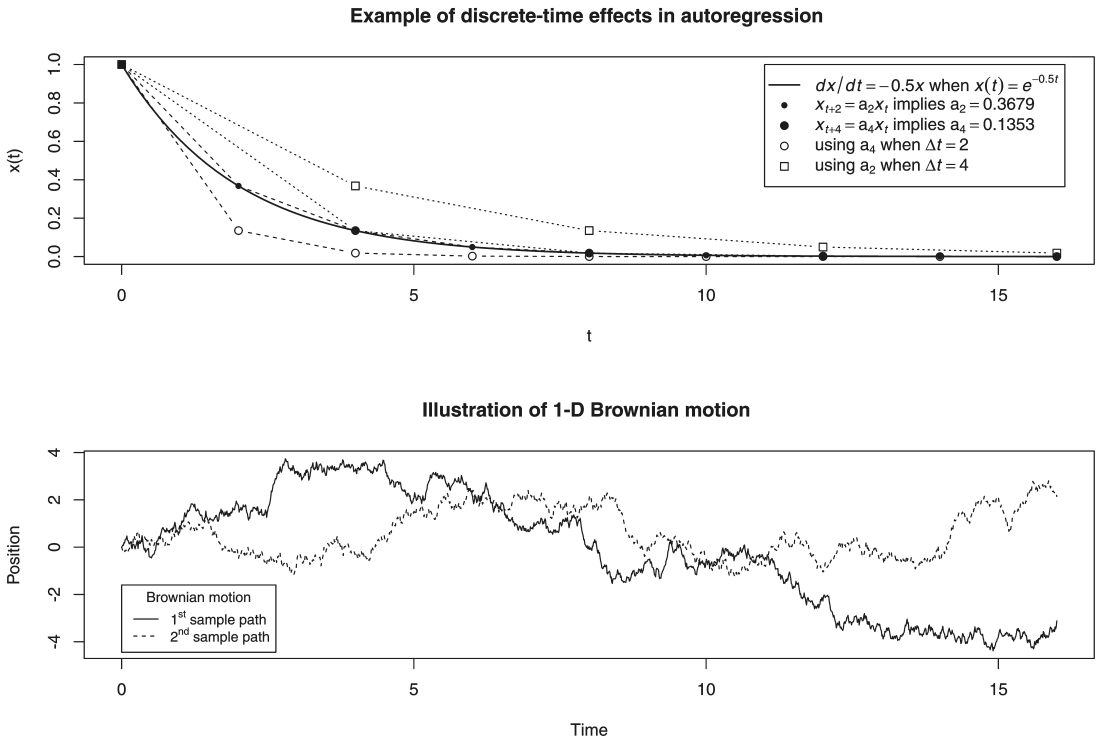

Figure 1. Illustration of discrete-time bias in a simple (non-random) autoregressive model and illustration of 1-D Brownian motion. First panel shows that a continuous-time linear differential equation model (solid line) can be accurately approximated with a discrete-time autoregressive model (solid circles), but these estimates do not apply to other time-intervals than the one used in their estimation (empty circles and squares illustrate application of an inconsistent interval). For simplicity, a deterministic situation was illustrated; similar problems also exist when modeling randomness, but are harder to detect. Second panel shows two example realizations, or paths, of the standard Wiener process. They are commonly known as 1-dimensional Brownian motions. The Wiener process is a model for the general properties of the paths (*e.g.*, ruggedness), allowing individual paths to differ from one other.

The Wiener process is only a part of the satisfactory model, however. In addition, the aim is to simultaneously model the stable, possibly heritable (Bienvenu et al., 2011; Cross-disorder Group of the Psychiatric Genomics Consortium, 2013a, 2013b; Sullivan et al., 2000), part of the between-individual variance, the age-declining deterministic trajectory of the population-average score (Galambos, Barker, & Krahn, 2006; Merikangas et al., 2003), state and measurement error variance, and the observation that perturbations due to exposures to life events tend to revert back to an individual-specific baseline (Clark, Diener, Georgellis, & Lucas, 2008; Kendler et al., 1998; Surtees & Wainwright, 1999). In technical terms, this thesis aims to alleviate the previous shortcomings in modeling of the depressive-symptoms sum score trajectories via combined use of stochastic differential equations (Øksendal, 2003) and structural equations (Bollen, 1989; Oud & Delsing, 2010). Only after that, the attention is turned to the individual symptoms, and to how such symptoms or behaviors might have emerged in evolutionary history.

1.3 Stable symptoms, symptoms associated with chronicity, and predominantly causally antecedent symptoms

If the contribution of depressive symptoms to the development of manifest depressive disorder varies as discussed above, an important practical question arises: do specific depressive symptoms predict the development of chronic depressive conditions particularly well? If this is the case, what are these symptoms? Are the symptoms that associate with a chronic course of depression also stable in time by themselves? If the symptoms form a causal network, as previously suggested (Cramer et al., 2012), are some symptoms predominantly causal antecedents rather than descendents for the other symptoms? These empirical questions have gotten little attention, probably due to the prevailing tradition to view the symptoms as reflecting a latent causal construct (Cramer et al., 2012, 2010). Here, the aim to avoid unnecessary assumptions without solid

empirical evidence implied the need to develop tools for studying the dynamics of the individual symptoms, in addition to their sum scores.

Because less modeling efforts for and empirical information about the symptom dynamics exists, the focus of this thesis is in contributing to the initiation of scientific inquiry into these issues, and in the provision of useful tools for future research. In addition to the above-mentioned questions of chronicity and stability, a limited excursion to inter-symptom causality estimation in the general population is provided. Both the experimental studies and the collection of temporally accurate longitudinal samples on depressive symptoms are difficult to carry out. Instead, many large cross-sectional data sets and data sets with low longitudinal sampling rates do exist. Therefore, an effective statistical method for inferring causality from cross-sectional samples would be a timely research tool.

As is well-known, statistical inferences cannot be made from cross-sectional *correlations*; a less known fact is that sometimes they can be made from other distributional properties than correlations (Hyvärinen & Smith, 2013; Hyvärinen, 2010; Shimizu, Hoyer, Hyvärinen, & Kerminen, 2006). Because these methods are novel and have caveats, caution is wise in their application. This thesis tests the methods on a benchmark data set, and applies them only to a single important causality problem, using several data sets and several depression inventories. The chosen single important causality problem was the predominant causality between sleep problems and other depressive symptoms.

Although sleep problems are included in the definitions and measures of depression, they often precede the onset of depressed mood and dysphoria (Almeida et al., 2011; Ford & Kamerow, 1989; Paunio et al., 2009; Perlis, Giles, Buysse, Tu, & Kupfer, 1997), and such temporality is considered to suggest causal antecedence (A. B. Hill, 1965). Temporality is not a definitive sign of causality, however, as sleep problems could be a prodromal symptom (A. B. Hill, 1965; Perlis et al., 1997), and also the opposite temporality is sometimes observed (Sivertsen et al., 2012). Furthermore, discrete-time cross-lagged models can lead to spurious causal inferences for dynamical systems in continuous time (Oud & Delsing, 2010). Hence, cross-sectional measures of causality

are a potentially important methodological addition to the epidemiologic research on causality among the depressive symptoms, especially if the symptoms form a causal network; fittingly, anecdotal examples of such networks have examined sleep problems (Cramer et al., 2012). Yet another reason for concentrating specifically on sleep-problems symptom versus the other depressive symptoms was that both sleep problems and depression are important public-health issues by themselves.

Complaints of poor sleep have been estimated to occur in up to 90% of the diagnosed cases of depression (Tsuno, Besset, & Ritchie, 2005), and poor sleep in itself can be a great burden for the individual and society. Sleep problems have been found to decrease work performance (Daley et al., 2009), increase risk of fatal and non-fatal accidents (Salminen et al., 2010; Åkerstedt, Fredlund, Gillberg, & Jansson, 2002), and to predict cause-specific work disability, physical and mental illnesses (Salo et al., 2010), and disability retirement (Lallukka, Haaramo, Lahelma, & Rahkonen, 2011). In addition to subjective complaints, also objectively measured short sleep duration is associated with poor cognitive performance and increased mortality (Fernandez-Mendoza et al., 2010; Vgontzas et al., 2010). Hence, regarding research on specific symptoms, this thesis concentrates on the issues of temporal stability, associations with chronic course of dysphoria, and pairwise causality between sleep problems and the other symptoms.

It is quite conceivable, however, that depressive symptoms will never be adequately understood from a 'proximate' viewpoint on the symptoms themselves. It may turn out to be necessary to adopt a more comprehensive evolutionary biological perspective in order to understand all the complexities in accumulated empirical evidence. As an analogy, consider fever, which may be an evolutionary adaptation. Fever associates with a multitude of peculiar circumstances, and is therefore difficult to understand when examined only as a host-specific disorder. Fever can be dangerous for the organism, even lethal; so, at first sight, it is also difficult to see fever as increasing the organism's evolutionary fitness (*i.e.*, fever as normal/evolved function). Once it is understood that most bacteria intruding a host organism cannot survive the temperature increase due to fever but viruses can (Chen & Shakhnovich, 2010), it becomes easier to see that fever can provide a sufficient average fitness increase to qualify as an adaptative response,

although it is not singularly beneficial and works only against a sub-population of pathogens. Therefore, the totality depends on a dynamic system that may transcend the boundaries of single host organisms, and cannot therefore be adequately understood by examining isolated host organisms. Same functions can incur both advantages and disadvantages, and it is the offset that matters. In an analogy, the final aim of this thesis is to provide a novel evolutionary biological perspective on the depressive symptoms.

1.4 Evolutionary origins of depressive symptoms

Depression is so commonly observed across the world that an increasing number of researchers have attempted to explain it from the perspective of evolutionary adaptations (Andrews & Thomson Jr, 2009; P. Gilbert, 2006; Nesse, 1991, 2000; Price, Gardner, & Erickson, 2004; Sloman, 2008; Watson & Andrews, 2002). The adaptive nature of severe depression has been questioned, however (Friedman, 2012; Nettle, 2004). It is difficult to see a prolonged complete cessation of action as an adaptive response to any situation: scientific adaptation arguments do not mesh with clinicians' intuition (Friedman, 2012), and they often suggest views that are incompatible with some subsets of the common depressive symptoms (Hagen, 2003, 2011). One of the more consistent evolution-related models is the *Bargaining Model* of depression that is concerned with social interactions among people; it proposes that one of the central functions of depression is to "withhold benefits from others until better terms are forthcoming" (Hagen, 2003). The lack of interest and inactivity (symptoms) that occur in depression deliver a "message" to a social group that for the depression sufferer there is little or no difference in fitness benefits obtained by investing heavily to current joint enterprises. If the depressed individual could provide something for the group, this "something" is withheld until better terms of cooperation are negotiated. Such a bargaining strategy can be effective when more direct coercion or switching of social groups are not.

The original Bargaining Model also draws from economic models of bargaining in the presence of private information (Hagen, 2003). Economic models often assume rationality and long-term planning, but in this thesis the driving force of strategic choices is behavioral imitation of presently successful individuals rather than rational long-term planning; as a means for robust evolutionary design, learning by imitation is to be favored over brain-based "generalized optimization engine" (Hagen & Hammerstein, 2006; Hammerstein, Hagen, Herz, & Herzel, 2006). The Bargaining Model argues that one function of the costly symptoms of depression is to "impose costs efficiently on other group members by withholding critical benefits, credibly signaling to them that one is suffering costs (Watson & Andrews, 2002), and compelling them to provide assistance or make changes" (Hagen, 2003). This core idea is embraced here as well, although the emphasis is on (unconscious) prompting of beneficial "changes" in one's collaborative environment rather than on unilateral direct "assistance". Table 1 summarizes the adaptive functions of DMS-IV symptoms in the Bargaining Model as originally proposed by Edward Hagen (2003), and these functions can be quite directly subsumed to the theoretical frame introduced in this thesis as well.

The Bargaining Model can explain symptoms that are difficult for other evolutionary accounts. For example, the analytical rumination hypothesis posits that depression is an adaptation that promotes cautious analysis of difficult, fitness-related problems (Andrews & Thomson Jr, 2009), but it is difficult to explain suicidal ideations from the viewpoint of adaptive re-evaluation of ones goals. From the viewpoint of the Bargaining Model, a completed suicide would most clearly remove an individual as a source of valuable benefits to others; a certain rate of suicide attempts is then necessary to underwrite the creditability of the threats, and the completed ones are a cost to maintain the creditability (Hagen, 2003). An evolution of suicide signaling/bargaining strategy is feasible when the rate of threats is much higher than the rate of attempts, and the latter is still much higher than the rate of completion; the average benefit received from influencing others over many generations by genes encoding for this strategy could exceed the cost suffered by the genes due to deaths by suicide (Hagen, 2003). Some adaptive theories have interpreted depression as an appeasement display designed to de-

escalate a social conflict situation (Price et al., 2004); yet, what appeasement is needed, say, in a case where depression derives from having a difficult child or a dying spouse? Instead, a difficult child or a dying spouse can be costly for a caregiver, and depression may serve to bargain the help of the wider social group (Hagen, 1999, 2003). More examples can be found elsewhere (Hagen, 1999, 2003, 2011; Watson & Andrews, 2002).

Although the depressive symptoms do reflect 'latent causes' of "perceived costs" and "compelling of others" in the Bargaining Model, it should be emphasized that they do not reflect a latent cause in the sense of a classical psychometric measurement model. Such latent-trait models assume that the latent causes directly and linearly modulate the items/symptoms that measure them, with the remaining residual variance in the symptoms being statistically independent of each other. Instead, the Bargaining Model only posits a unifying goal, and the individual symptoms serve different functions for achieving the desired goal (Table 1), their presence being conditional on environmental conditions beyond those that originally motivated the goal. Therefore, the symptoms are not assumed conditionally independent given a latent cause(s); that is, the principle of "local independence" should not hold in this important special case, as generally suspected by several measure theoreticians (Borsboom et al., 2003; Reise & Waller, 2009).

Although much has been written about depression as an individual-level adaptation, less material exists about the meaning of depression for the adaptive fitness of a group of individuals, or society; even though depressive symptoms have been shown to spread in social networks (A. L. Hill, Rand, Nowak, & Christakis, 2010; Rosenquist, Fowler, & Christakis, 2010). This thesis aims to outline the implications of the Bargaining Model for the emergence of cooperation under social-imitation dynamics (Nowak, 2006, 2012; Sigmund, 2010) and for the resulting fitness of groups of individuals, thereby providing an evolutionary sociobiological model for depressive-symptom dynamics. It is argued that by combining the Bargaining Model with mathematical models for evolution of cooperation, a more comprehensive theory for adaptive origins of depressive symptoms is achieved.

Table 1. Symptoms of Major Depressive Episode according to the DMS-IV, and Their Adaptive Interpretations according to the Bargaining Model (Hagen, 2003)

Symptoms of Major Depressive Disorder	**Hypothesized Functions according to the Bargaining Model**
1. Sad or depressed affect	Information to the sufferer that the current social strategy or circumstance is imposing a net fitness cost
2. Marked loss of interest in virtually all activities	a) Reduce investment in the costly strategy (minor depression) b) Reduce investment in oneself and others (major depression)
3. Significant weight loss or gain	Loss: reduce investment in oneself (Gain: store resources for tough times ahead. Weight gain was probably difficult in environment of evolutionary adaptedness)
4. Hypersomnia or insomnia	Hypersomnia: reduce productivity (Insomnia: allocate additional cognitive resources toward finding a profitable resolution to the current crisis)
5. Psychomotor retardation or agitation	Retardation: reduce productivity (Agitation: comorbid anxiety. Conflicts with social partners are often dangerous)
6. Fatigue or loss of energy	Reduce productivity
7. Feelings of worthlessness or guilt	Worthlessness: contributions undervalued by others Guilt: defecting from social contracts imposes costs on others
8. Diminished ability to think or concentrate	Reduce productivity (and, more importantly, divert cognitive resources to renegotiating the current venture or toward finding more profitable alternatives)
9. Recurrent thoughts of death	Threaten to put future productivity at risk

Note: the table is reproduced with permission from Hagen, 2003, page 101

An option to abstain from a joint enterprise may be critically important for the emergence of cooperation in joint enterprises of many self-interested agents (De Silva, Hauert, Traulsen, & Sigmund, 2010; Fowler, 2005; Hauert, Traulsen, Brandt, Nowak, & Sigmund, 2007; Sigmund, De Silva, Traulsen, & Hauert, 2010; Sigmund, 2010). The Bargaining Model (Hagen, 2003) and similar accounts suggest that depression offers a way to not participate in a joint enterprise while, at the same time, retaining important

baseline benefits. Therefore, the combined hypothesis (detailed in Methods section) is called the *nonparticipation hypothesis of depression*, in short, the nonparticipation hypothesis. The main contribution of this new hypothesis for the adaptive origins of depression is that it mechanistically connects genetic, environmental, and dynamic social effects. Therefore, it contributes to the complex task of determining "which depressions are adaptive in which context" (P. Gilbert, 2006). Importantly, it is not implied that depression would necessarily be adaptive for a single individual suffering from it, or for a single loved one of a depressed person. The argument implies that because of the naturally occurring social dynamics, it is adaptive for a group of individuals that some or all have a *capacity* for depression.

The nonparticipation hypothesis is not in direct conflict with existing proximal individual-based theories of depression, but can co-exist with them, providing a higher-level sociobiological theory. The non-participation hypothesis aims to explain a diverse set of seemingly unrelated empirical observations, including comorbidity among negative emotions (Merikangas et al., 2003), cascading of emotions in social networks (A. L. Hill et al., 2010; Rosenquist et al., 2010), connection between income inequality and depression (Fiscella & Franks, 2000; Godoy et al., 2006; Muramatsu, 2003; Weich, Lewis, & Jenkins, 2001), and non-specific benefits of the therapeutic alliance (Martin, Garske, & Davis, 2000).

2 AIMS OF THE STUDY

The first aim of this study is to find a better statistical model for longitudinal trajectories of depressive-symptoms sum scores than the previous ones. This model should take into account that (a) the sum scores are "pushed and pulled" by many factors, (b) the individual-level perturbations tend to revert back to an individual-specific baseline, (c) a declining age trend exists in populations' average score, and that (d) observed people are more abundant in existing epidemiological data sets than observed temporally distinct samples per people. Further important points are that the modeling approach

needs to cope not only with sparse temporal sampling but also with (e) varying sampling intervals and (f) partially missing data that are commonplace in large epidemiological samples. A suitable statistical model has been recently introduced in the methodological literature (Oud & Delsing, 2010; Oud & Jansen, 2000), and will be applied in a large general-population sample with a total of 16-year follow-up and four prospective samples from the same individuals (Raitakari et al., 2008). A second, and related, goal is to compute from the estimated statistical model the relative contributions of (1) stable trait, (2) accumulating effects, and (3) pure state variations in the total variance of depressive-symptoms scores.

The Introduction discussed that the individual depressive symptoms forming the sum scores and diagnoses can be differentially perturbed by different events, and by other depressive symptoms. Although in stochastic models these effects can be encapsulated into random movements of the sum scores, it is also possible to further analyze them. Due to the practical needs and historical emphasis on the latent-cause formulations, the study of sum scores and diagnoses has often preceded the systematic study of their constituent parts, the symptoms themselves. Also, as the study of symptoms involves high-dimensional system(s), the task is challenging and must be tackled in manageable pieces. This thesis starts symptom-level analyses by studying two obviously important questions about the temporal development of individual symptoms. Specifically, the third aim of the thesis is to determine which symptoms are most stable in their levels through time. The fourth, and related, aim is to explore which depressive symptoms, if any, are associated with chronic trajectories of dysphoria; that is, with a chronically elevated depression sum score. A high score in a stable symptom naturally predisposes for a stably high sum score, as the symptoms are positively correlated with the sum score.

The fifth aim of this thesis is to introduce statistical methods of pairwise causality estimation to the depressive-symptoms research; a worthy additional tool should the depressive symptoms form a causal network (Borsboom, Cramer, Schmittmann, Epskamp, & Waldorp, 2011; Cramer et al., 2012, 2010). The sixth, related, aim is to apply these methods in order to estimate predominant causality between the sleep-

problems symptom and other depressive symptoms in the general population. This chosen symptom has the advantage of connecting with an important and much studied public-health topic.

In addition to the intermediate-level of analysis involving sum scores and diagnoses and the fine-grained analysis of depressive symptoms, the Introduction section noted that a holistic evolutionary analysis may be necessary for a deeper understanding of depression. Thus, the seventh and the final aim of this thesis is to connect depressive symptoms with the evolutionary history through the "nonparticipation hypothesis". Overall, this thesis aims to apply multiple levels of analysis in order to better understand depression.

3 METHODS

3.1 Cardiovascular Risk in Young Finns study

Data from two separate population studies were used. First, the Cardiovascular Risk in Young Finns study, is an on-going population-based cohort study (Raitakari et al., 2008; Åkerblom et al., 1991); its participants have provided a written informed consent, and it has been approved by the ethical committee of the Varsinais-Suomi's hospital district's federation of municipalities.

3.1.1 Participants

The original Young Finns sample consists of 3596 healthy Finnish children and adolescents derived from six birth cohorts, aged 3, 6, 9, 12, 15, and 18 years at the baseline year 1980. In order to select a broadly sociodemographically representative sample, Finland was divided into five areas according to locations of university cities with a medical school (Helsinki, Kuopio, Oulu, Tampere, and Turku). In each area, urban and rural boys and girls were randomly selected on the basis of their unique

personal social-security number. The sample has been followed subsequently in 7 data collection waves in 1983, 1986, 1989, 1992, 1997, 2001, and 2007-08; practically all participants represented white European origin. A detailed description of the cohort can be found from earlier cohort-profile publications (Raitakari et al., 2008; Åkerblom et al., 1991). The follow-up in 2007-08 included questions on sleep and socioeconomic status, as well as two different questionnaires on depression. Data from 1992, 1997, 2001, and 2008 contained follow-up of a same depression inventory, and are therefore used in time series analyses of depressive symptoms. In addition, 1348 participants' parents' socioeconomic status (SES) in the year 1983 was used in a benchmark analysis of the pairwise causality statistics.

3.1.2 Measures

Altogether three different depression inventories were applied in this thesis. Two of them were measured in the Young Finns data set. The Young Finns study included a modified version of Beck's Depression Inventory (mBDI) that has been measured four times. In addition, the latest follow-up included Beck's Depression Inventory II [BDI-II (Beck, Steer, & Brown, 1996)]. Originally Beck described four phenotypes for each depressive symptom that corresponded to not present, mild, moderate, and severe forms of the underlying symptom, and he developed questionnaire items for each phenotype; the scoring system yielded a weight 0 through 3 points for the none through severe symptom phenotype (Beck & Steer, 1993; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961; Beck, 1967). The mBDI represents the mild phenotypes of the original items of Beck's Depression Inventory with a five-point scale ranging from 1 = "description does not fit me" gradually to an explicit 5 = "true". This scoring system has been thought to involve less threatening material [*e.g.*, it avoids confronting statements like "I would kill myself if I could" (Beck et al., 1961)], to better capture the subtler general-population variance compared to clinical tools, and to be less time-consuming and easier to fill out compared to the original inventory (Katainen, Räikkönen, &

Keltikangas-Järvinen, 1999; Katainen, Räikkönen, Keskivaara, & Keltikangas-Järvinen, 1999); although, explicit empirical evidence behind these rationale seems to be lacking thus far. The BDI-II in the Young Finns study retained the Beck's original phenotype-scoring system.

Both mBDI and BDI-II included 21 items (range 1-5 and 0-3, respectively), and sums of the items represents mBDI and BDI-II scores. In addition to sums, item averages, *z*-score transformations of the sums (translate average to zero and scale variance to one), and affine transformations of the sums (some other translation and scaling) were assessed. When relationships between depression and sleep problems were assessed, sleep-related items were removed from the depression inventories; however, a sensitivity analysis indicated that results were almost the same also without this precaution. Due to space limitations, sensitivity-analysis results are not shown here, but can be found from a supplement of the original contribution III.

Parents' socioeconomic status (SES) was defined in the Young Finns data as follows. Total income of the family was inquired from the parents in the year 1983, when the participants (offspring) themselves were 6-21 years old; income prior to taxes (but after the removal of investments necessary for the obtaining of income) was recorder with an eight point precision. Number of years of education until the year 1986 was recorded for mother and for father; participants lacking one or both parents were excluded from the analysis, as were those lacking any of the required data. The level of education was also inquired from both parents with a seven point precision: 1 = "elementary school unfinished"; 2 = "elementary school"; 3 = "middle school/part of high school"; 4 = "high school"; 5 = "college-level"; 6 = "vocational/technical/business school"; 7 = "university". All the five variables were (z-score) standardized, summed, and standardized again, in order to form the parents' SES variable.

Offspring's SES was defined as follows. Gross income of the participant in the year 2007 was inquired with an eight point precision. Years of education that a participant had undertaken until the year 2007 was recorded; the age of participants ranged from 30 to 45 years at that point. Also the level of education was recorded with an eight point precision: 1 = "vocational school or any kind of course or institute degree"; 2 =

"technical college", 3 = "vocational high school"; 4 = "university studies without degree"; 5 = "bachelor's degree"; 6 = "master's degree"; 7 = "licentiate"; 8 = "doctor". All the three variables were (z-score) standardized, summed, and standardized again, in order to form the offspring's SES variable.

In the Young Finns data, sleep problems were assessed with the Jenkin's scale consisting of four items that assess: difficulties falling sleep, frequent awakenings, troubles staying asleep (including too early waking), and feelings of tiredness and exhaustion after a regular night of sleep (Jenkins, Stanton, Niemcryk, & Rose, 1988). These items were answered with the following six-point precision: 1 = "not at all"; 2 = "1-3 nights in month"; 3 = "1 night in week"; 4 = "2-4 nights in week"; 5 = "5-6 nights in week"; and 6 = "every night"; the average of the four items formed the final measure of Sleep problems. Cronbach's internal consistency (alpha) for Sleep problems was 0.77.

3.2 Wisconsin Longitudinal Study

As a second data set, a part of the Wisconsin Longitudinal Study data was used. The data were initially collected via a telephone interview, after which a questionnaire was mailed to the participants (Hauser, Sewell, & Hauser, 2006). Informed consent was obtained at the beginning of the telephone interview. All instruments and operations were approved by the Institutional Review Board of the University of Wisconsin-Madison.

3.2.1 Participants

The Wisconsin Longitudinal Study (www.ssc.wisc.edu/wlsresearch) is a prospective cohort study of a random sample of 10 317 participants (5326 women, 4991 men) born between 1937 and 1940, and followed since they graduated from Wisconsin high

schools in 1957. After baseline data collection in 1957, survey responses were collected in 1964, 1975, 1992-93, and 2004. The sample is broadly representative of white, non-Hispanic US men and women who completed at least high school education. It is estimated that about 75% of Wisconsin youth graduated from high school in the late 1950s – everyone in the primary sample graduated from high school. A mail questionnaire collected in 1992-93 contained a depression inventory and 3 sleep items (Hauser et al., 2006); only these parts of the Wisconsin data were used in this thesis.

3.2.2 Measures

In addition to the Young-Finns measures originating from Beck's work, another depression inventory was applied in the Wisconsin Longitudinal Study; this was a modified version of the Center for Epidemiologic Studies Depression scale (mCES-D). The measure is practically the same as the original CES-D (Radloff, 1977), only a slightly more accurate seven-point scoring was applied to the items that enquire on how many days during the past week the participant suffered from a given symptom. The mCES-D for a participant was a sum of 20 items, each of whose scores could range from 0 to 7 (total range 0-140); this sum score, described and computed by the Wisconsin researchers (Hauser et al., 2006), was applied in here too.

In the Wisconsin data, sleep problems were coded with zero if the participant had answered that he or she did not have trouble sleeping in the past six months. Otherwise, they were coded as a sum of two items with the following content: "How often have you had trouble sleeping?" (1 = "monthly or less often"; 2 = "about once a week"; 3 = "daily or more often") and "How much discomfort has trouble sleeping caused you in the past six months?" (0 = "none"; 1 = "a little"; 2 = "some"; 3 = "a lot").

3.3 Public good games and nonparticipation

In addition to the above-described empirical measures, this thesis investigates a possible theoretical model for depression. For this, some definitions are needed. Evolutionary adaptations have been defined as "inherited and reliably developing characteristics that came into existence through natural selection because they helped to solve problems of survival or reproduction during the period of their evolution" (Buss, 2008). The underlying assumption here is that, during history, cooperation among bands of humans has solved problems of survival. This is a notion that many concur with, starting from Darwin: "There can be no doubt that tribe including many members who … were always ready to give aid to each other and sacrifice themselves for the common good, would be victorious over other tribes; and this would be natural selection" (Darwin, 1871; Nowak, 2012; Sigmund, 2010).

In addition to intuition, both historical observations (Bowles, 2009) and experimental studies (Puurtinen & Mappes, 2009; Sääksvuori et al., 2011) suggest that between-group competition facilitates altruistic behavior and cooperation within group, with more cooperative groups faring better in the between-group competition. Recently, a relatively simple model based on an interplay of individually costly group formation/cooperation and introduction of new military techniques was able to explain 65% of the variance in spatiotemporal distribution of historical large-scale societies in Afroeurasia between 1500 BCE and 1500 CE, suggesting that warfare between human groups has been a strong selective force (Turchin, Currie, Turner, & Gavrilets, 2013).

Yet, natural selection cares not of the intent, but of the outcome; cooperation arising from immediate selfish interests can be as efficient an adaptation as cooperation due to other intentions, and may be easier to establish among strangers than self-sacrifice. Striving for the simplest dynamic model to begin with, this theoretical study is concerned with selfish motivations as a driving force of social dynamics [*e.g.*, as opposed to kin-selection (Nowak, 2012; Sigmund, 2010)].

3.3.1 A model for joint enterprise

When attempting to explain the emergence of cooperative behavior in joint enterprises, it is customary to begin with agents/individuals who attempt to optimize some selfish gain, and then derive the conditions where the selfish effort leads to cooperation among the agents (Axelrod & Hamilton, 1981; Nowak, 2006; Sigmund, 2010). A Public Good Game consists of several rounds of interaction among subsets of individuals from a population of M individuals [a well-behaved infinite-population equivalent is not available for the particular four-strategy game that is of interest here (Hauert et al., 2007; Sigmund, 2010)]. In each round, N players are randomly drawn from the population of M individuals and given an option to participate in a joint enterprise, where they can choose to invest a cost, or effort, c. The efforts are then multiplied by the factor r ($r > 1$), yielding the benefit rc that is evenly distributed among those participating in the enterprise—regardless of who paid the costs. After each round, there is an option to pay more in order to impose a punishment for those who participated, reaped the benefits, but did not contribute the payment c for the public good.

The population of M individuals consists of players applying one of the four alternative strategies: M_c Cooperators pay the cost in each round; M_d Defectors (the free-riders) never pay; M_p Punishers pay the cost in each round, and also pay γ per each Defector in the enterprise in order to impose an unavoidable fine of βM_p for each of the Defectors (where $\beta > \gamma$); finally, M_n Nonparticipants do not take part in the game, but gain an amount σ per round independently of the game (if N-1 Nonparticipants happen to be selected, the enterprise is canceled from that round, and the remaining participant also gains σ). The four strategies are the only alternatives: $M_c + M_d + M_p + M_n = M$. For the game to be both potentially lucrative and such that nonparticipation is a real option, inequality $0 < \sigma < (r - 1)c$ must hold; that is, when all players cooperate it is advantageous to participate, but when all defect a Nonparticipant does better. For non-triviality, the game cannot be beneficial regardless of other players actions (*i.e.*, $cr/N - c < \sigma$ must hold).

Models for Public Good Games usually apply fitness-based *imitation dynamics*: successful strategies are imitated by other players in the population with an intensity proportional to the product of relative fitness of strategy and the frequency of the players using that strategy (Sigmund, 2010). The frequency of a given strategy in the population can be interpreted as the frequency for observing the associated behavioral model, which is directly related to occasions where "imitation" of that strategy can occur. Because the fitness of a strategy usually depends on the distribution of strategies among the players, fitnesses of strategies change with the proportion of participants using them. Therefore, imitation can give rise to rich temporal dynamics. In both evolutionary and imitation dynamics, it is typically assumed that rare 'innovations' exist; that is, a randomly chosen strategy is introduced on rare occasions via behavioral exploration or mutation (Nowak, 2006; Sigmund, 2010). A small rate of innovations is assumed here too.

Some examples of how proportions of played strategies affect their fitness are apparent: If most co-players are Defectors, it is wisest to be a Nonparticipant and gain σ. A solitary Punisher would need to punish left and right, imposing huge costs for the self. A Defector would gain nothing, and a Cooperator would lose $c(1 - r/N)$. Therefore, imitation dynamics will favor Nonparticipants. But as they take over the population, the game becomes most lucrative for Cooperators. Again, when Cooperators are a majority, the incentive to defect is obvious: the gain is proportional to rc/N compared to the Cooperators' $rc/N - c$. This type of oscillation of three strategies is known as "rock-paper-scissors" dynamics (Nowak, 2006; Sigmund, 2010); in very large population, no one then benefits more than the Nonparticipants' σ in the long run [the time-average of the pay-offs is σ (Sigmund, 2010)]. If participation in the game is compulsory, that is, only Cooperators, Defectors, and (costly) Punishers play, then Defectors dominate the time-average of the strategy-distribution (De Silva et al., 2010; Hauert et al., 2007; Sigmund, 2010); the average pay-off in the population is less than σ. The 'tragedy of commons', evaporation of mutually beneficial cooperation, occurs. Therefore, the game cannot be simplified without compromising the adaptive value of the joint enterprise for

the population. Only when all the four strategies are present, the collaborative effort flourishes most of the time (De Silva et al., 2010; Hauert et al., 2007; Sigmund, 2010).

In principle, one could conceive many more possible strategies and game-settings; however, the above model is simple and its predictions show remarkable robustness for alternative formulations (De Silva et al., 2010; Sigmund et al., 2010; Sigmund, 2010), a key factor in evolutionary design (Hammerstein, Hagen, Herz, & Herzel, 2006). For example, punishing can be implemented via "sanctioning institution" supported by the Punishers instead of the above-defined "peer-punishing" without changing the essential result regarding the importance of the voluntary participation (Sigmund et al., 2010). The option to abstain from the joint enterprise is necessary for the coerced cooperation to emerge in most model formulations. For this reason, and because it is so difficult to find simple dynamical settings that encourage cooperation over defection, the discussed four-strategy model is most likely to have wide-ranging biological and social relevance. The next section outlines how depression could be interpreted as nonparticipation, and why this might have been crucial during the evolution of the cooperative human.

3.3.2 Nonparticipation via depressive symptoms

Throughout the evolutionary past, humans have been highly dependent on cooperative efforts (Buss, 2008). For example, in small bands of hunter-gatherers it has been adaptive to share the uncertainty in food acquiring. The temptation to defect has obviously existed in the past, and may have even resulted in genetic adaptations (Buss, 2008). Also aggression and punishing behaviors have certainly existed in the past. However, in close-knit communities, the option to not participate may sometimes have been less evident. Assuming that a solitary human faces difficulties in surviving in the wild, the option to not 'play' cannot be taken as granted (MacDonald & Leary, 2005). In the above section, however, it was discussed that the option to abstain is crucial for the emergence of constructive cooperation.

Depression may have offered a means to abstain from a joint enterprise while simultaneously avoiding enforcement from others. Depressed persons suffer from fatigue, inefficiency, and lack of effort. Even brief interactions with a depressed person suffice to convince others that they are not going to get the kind of relationship they want (Boswell & Murray, 1981; Coyne, 1976). Normally this would run a risk of punishment or ostracism by the group, but the depressed compensate by eliciting empathy in others: "The symptoms of depressed person are aversive yet powerful in the ability to arouse guilt in others and to inhibit any direct expression of annoyance and hostility from others" (Coyne, 1976) and people are less likely to retaliate against depressed than cheerful person (Surbey & Simpson, 2010). Depression may thus be an effective way to not participate and still maintain modest benefits (the Nonparticipants' σ). Choosing the modest Nonparticipants' share is bargaining in the precise sense of the Bargaining Model; that is, one withdraws his share of effort from the 'Defectors' until more cooperative terms of play emerge (Hagen, 2003). Even though bargaining were not a conscious act and no direct benefit would come for the depressed person nor any other single individual, depression as a phenomenon may have benefited groups of individuals engaging in joint enterprises by promoting the emergence of cooperation in social dynamics.

Depressed mood can instantiate a nonparticipation strategy more effectively than many other adverse conditions. For example, schizophrenia also elicits aversive reactions in healthy people (Boswell & Murray, 1981), but a schizophrenic does not readily switch strategies and start fully re-participating in a joint enterprise; it is not as dynamic condition as depression. Depression and low moods preserve potential for recovery/change, as required in imitation dynamics: more than half of the patients remit spontaneously within a year according to some estimates (Whiteford et al., 2012). In addition to withdrawal, depressed mood is thought to signal submission in a conflict situation; that is, it carries a behavioral message "I will not retaliate your actions" (Price et al., 2004). Therefore, the depressed are not easily mistaken by others as applying a punishing strategy, despite their obvious dissatisfaction regarding a state of affairs. Instead, schizophrenia (causing, *e.g.*, blunted affect) does not appear designed for

eliciting the empathy of others, and therefore it lacks an obvious mechanism that would guarantee any Nonparticipants' pay-off. Involved delusions and disorganized behavior may rather cause social conflicts than aid in avoiding them. The broad ideas discussed in this section will be placed into an explicit model in section 3.4.4.

3.4 Statistical and mathematical models

3.4.1 Depression-inventory distributions

Before proceeding to other analyses, the distributions of the depression inventories (*i.e.*, of sum scores) were assessed in various ways. First, standard summaries are presented. Second, the default (Gaussian) kernel density estimate ("density" function in "stats" package) of the statistical R-software was applied to the sum score variables [software version 2.15.3 in 64-bit Linux was always used (R Core Team, 2012)]; a density is like continuous 'histogram' that integrates/sums to one. Third, the relationship between mBDI-II and BDI-II was assessed via a scatterplot and classical (Ordinary Least Squares) regression methods; relationships between mCES-D and the other scales could not be assessed, as measures from overlapping participants were lacking. Fourth, the mBDI and BDI-II were compared in a same scale by estimating a Graded Response Model with a logit-response function simultaneously to items of the both scales ["ltm" package version 0.9-9 (Rizopoulos, 2006; Samejima, 1997)]; the model was used in an approximative sense, as the latent unidimensionality was not expected to hold strictly (see section 1.1). Fifth, the relationship between depressive-symptom sum scores and the external sleep-problems variable was assessed by scatterplots and regression.

In general, deviations from Normal/Gaussian distribution were tested using Lilliefors test (Lilliefors, 1967), and the magnitude of the deviations was mainly assessed from kurtosis and skewness of the distribution. 'Effect sizes' of a binary grouping on averages was sometimes computed as Cohen's $d = (m_1 - m_2)/s$, where m_1 is the average value in first group and m_2 in the second group, while s is the pooled standard

deviation; a crude rule of thumb is that 0.2 to 0.3 is small effect, around 0.5 is a medium effect, and 0.8 or more is a large effect (Cohen, 1988).

3.4.2 Stochastic model for sum-score trajectories

The model that was applied to the longitudinal mBDI sum-score data herein corresponds to a *continuous* version of the autoregressive Latent Trait-State Model (Cole & Martin, 2005; Kenny & Zautra, 1995). The difference in the chosen approach compared to Latent Trait-State Model is that one models a time-continuous process $x(t)$ instead of just the four discrete successive latent variables that correspond to the theoretically arbitrary measurement times t_1, t_2, t_3, and t_4, (here, the years 1992, 1997, 2001, and 2008). The model for $x(t)$ is more widely known as the Linear Stochastic Differential Equation Model (Oud & Delsing, 2010; Øksendal, 2003), but it is estimated here using a Structural Equation Model with appropriate nonlinear constraints (Oud & Delsing, 2010; Oud & Jansen, 2000). In that case, the model has been also referred to as the Exact Discrete Model (EDM), because it is in exact equivalence with the underlying continuous time process (Oud & Delsing, 2010). Due to technical demands, continuity is rarely modeled in practice, although it has been recognized for a long time that the discrete approximation is problematic (Bartlett, 1946). Some of the associated problems were illustrated in Figure 1, and more discussion can be found from previous methodological research (Oud & Delsing, 2010).

Let us denote the depression score at time t for participant i by random variable $x_i(t)$, whose realizations involve fixing the time and the participant. Using existing notation (Oud & Delsing, 2010), the description of the time-evolution that is fitted to the data is of the form

$$dx_i(t) = Ax_i(t)dt + bdt + \kappa_i dt + GdW_i(t) \qquad (1)$$

that reads aloud as follows. Infinitesimal, or instantaneous, change in the process $x_i(t)$ depends on the multiple of the value of the process itself at the time t by a constant parameter A, on the value of the constant parameter b, on the time-constant between-subjects random variable κ, and on the multiple of the instantaneous change in the Wiener process W by the constant parameter G at the time t. Hence, the model implies a participant-specific trait level, controlled by the 'random-effect' parameter κ_i, and participant-specific time-accumulating random perturbations due to increments in Brownian paths modeled by $GdW_i(t)/dt$. In addition, a population-level trend controlled by A and b is estimated, as well as population-level magnitude of time-accumulating random perturbations, G. It is also assumed that, at any time t, for any participant i, the recorded observations of the process $x_i(t)$ are masked by measurement error or state variations unrelated to the time evolution of the process; that is, any observation $y_i(t)$ is of the form

$$y_i(t) = x_i(t) + \varepsilon_i(t), \tag{2}$$

where $\varepsilon_i(t)$ is a Normally distributed random variable with mean zero and variance Θ, independently and identically for all different individuals and all different measurement times. Naturally, $W_i(t)$, κ_i, $\varepsilon_i(t)$, and $\varepsilon_i(s)$ are also independent of each other at all times t and s. The aim is to estimate the parameters of the model, including parameters of the assumed Normal distribution for initial time-point $x_i(t_0)$, and to use the estimated model in order to construct a new kind of decomposition of observed depression variance into trait, cumulative, and state components.

The solution of the stochastic differential equation 1 is

$$x_i(t) = e^{A(t-t_0)}x_i(t_0) + A^{-1}\left(e^{A(t-t_0)} - 1\right)(b + \kappa_i) + \int_{t_0}^{t} e^{A(t-s)}GdW_i(s), \tag{3}$$

where e is the natural base of exponential function, and t_0 is the time of initial measurement; time can be translated so that t_0 is zero without a loss in generality. The

last summand in the right side of the equation is an instance of Itô integral, according to the Japanese mathematician Kiyoshi Itô (Klenke, 2008; Øksendal, 2003). For the Itô integrals of integrable functions f and g, the following general rules of expectations apply:

$$E\left[\int_{t_0}^{t} f(s)\,dW(s)\right]=0,$$

and

$$E\left[\int_{t_0}^{t} f(s)\,dW(s)\int_{t_0}^{h} g(s)\,dW(s)\right]=E\left[\int_{t_0}^{\min(t,h)} f(s)\,g(s)\,ds\right].$$

As shown by Equation 3, at any time t, $x_i(t)$ is a sum of three independent random variables, respectively associated with $x_i(t_0)$, κ_i, and the integral with respect to the Wiener process (b is constant). Hence, the rules of Itô-integration and elementary rules of expectations in probability calculus (Klenke, 2008) can be applied to the Equations 2 and 3 in order to compute the population variance of the measurements $y(t)$.

Algebraic manipulation of the variance term, $\mathrm{Var}[y(t)] = \mathrm{E}[y(t)^2] - \mathrm{E}[y(t)]^2$, leads to an expression that can be written as a sum of four terms related to initial depression value, trait variation, variance due to cumulative effects, and variance due to state effects or measurement errors. More specifically, $\mathrm{Var}[y(t)] = V_{init}(t) + V_{trait}(t) + V_{cum}(t) + V_{state}(t)$, wherein

$$V_{init}(t)=e^{2A(t-t_0)}\Phi_{x(0)}+2\,e^{A(t-t_0)}A^{-1}\left(e^{A(t-t_0)}-1\right)\Phi_{\kappa,x(0)},$$
$$V_{trait}(t)=A^{-2}\left(e^{A(t-t_0)}-1\right)^2\Phi_{\kappa},$$
$$V_{cum}(t)=(2\mathrm{A})^{-1}\left(e^{2\mathrm{A}(t-t_0)}-1\right)G^2, \text{ and}$$
$$V_{state}(t)=\Theta.$$

In these expressions, some further estimable parameters are introduced; namely, $\Phi_{x(0)}$ is the population variance of the initial value (in year 1992) of process $x(t)$, $\Phi_{\kappa,x(0)}$ is the population covariance between initial values and trait variable, and Φ_{κ} is the population variance of the trait variable. In a stable process that does not increase or decrease without bounds, A is negative, implying that all the variance components given by the above equations are positive at all times. Then the initial-value contribution to variance vanishes as time passes (*i.e.*, the terms that depend on $x(0)$—shorthand for $x(t_0)$ in sub-indices).

In the infinite-time limit, $\lim_{t\to\infty}\mathrm{Var}[y(t)] = V_{trait} + V_{cum} + V_{state}$, where the stable trait variance is $V_{trait} = A^{-2}\Phi_{\kappa}$, variance due to cumulative effects is $V_{cum} = (2A)^{-1}G^2$, and variance due to the state/measurement error is $V_{state} = \Theta$; the latter component relates to difference between the latent process of equation 3 and its measurement (equation 2). In the results section, the three-component variance decomposition of the measured depression score is presented for the infinite-time limit. While this is one way to interpret the course of depression without confounding due to variation in initial measurements, the growing proportion of variance attributed to cumulative effects is also plotted as a function of time, that is, the graph is provided for the function $f(t) = V_{cum}(t)/\{V_{init}(t) + V_{trait}(t) + V_{cum}(t) + V_{state}(t)\}$. This serves to estimate how fast new cumulative effects override the ambiguous initial-point variance that mixes both cumulative and trait sources of variance (the original contribution I studied an analogous f with V_{init} removed from the denominator).

Before the variance components were computed, parameters of the model (*i.e.*, EDM) were estimated, and negativity of the parameter A verified. Here, the EDM was estimated using OpenMx structural equation modeling software (Boker et al., 2011), but also other structural-equation software that implements non-linearly constrained estimation can do the job. In estimation of a structural equation model, one matches the theoretical mean and covariance matrices to the empirically observed ones by algorithmically adjusting the free parameters of the model. From equations 2 and 3, one

sees that, when setting $t_0 = 0$, the covariance between any two measurement times t and s having a minimum min(t, s), is

$$Cov[y(t), y(s)] = e^{A(t+s)}\Phi_{x(0)} + \left(e^{At}A^{-1}(e^{As}-1) + e^{As}A^{-1}(e^{At}-1)\right)\Phi_{x(0),\kappa} + A^{-2}(e^{At}-1)(e^{As}-1)\Phi_{\kappa} + (2\mathrm{A})^{-1}\left(e^{A(t+s)}-e^{A(t+s-\min(t,s))}\right)G^2 + 1_{\{0\}}(t-s)\Theta,$$

and the expectation is just $E[y(t)] = e^{At}\mu_{x(0)} + A^{-1}(e^{At} - 1)b$, where $\mu_{x(0)}$ is an estimable parameter describing the expected population mean at the initial zero-time measurement ($1_{\{0\}}$ is just the indicator function, getting the value 1 when $t = s$, and 0 otherwise). Altogether the model introduces eight free parameters (listed in Abbreviations and Symbols -section) to be estimated using the 14 available degrees of freedom in the empirically observed means and covariances of the four successive Young Finns measurements. Although there are many ways to enter covariance matrices into structural equation software, it is crucial in EDM that the above nonlinear dependencies on time are correctly worked into the expected covariance matrix. Perhaps the most detailed example is provided by Oud and Delsing (2010), but several sources provide further practical instructions (Oud & Jansen, 2000; Oud & Singer, 2008).

The structural equation modeling, or EDM, approach tends to yield similar results as the more traditional Kalman-filtering approaches to estimation (Oud & Singer, 2008), but provides a useful suite of additional tools, such as a tested and efficient algorithm for handling the missing data. More specifically, the model was estimated with Full Information Maximum Likelihood (FIML) approach that has been found successful in the modeling of missing data (Muthén, Kaplan, & Hollis, 1987), even in the case of non-normal data (Enders, 2001), and has been implemented to the OpenMx software. In addition to missing-data handling, the structural equation modeling framework provides standard tools for assessing model fit and for the comparisons among models.

Structural equation models were evaluated using Likelikood-ratio test, and by Akaike's (AIC) and Bayesian (BIC) information criteria (Bollen, 1989; Kass & Raftery, 1995). A model that attains lower AIC (respectively, BIC), value than the competing

models is the preferred model according to AIC (respectively, BIC). Bootstrapping with 1000 bootstrap re-samples of original data was used in order to derive percentile confidence intervals for the EDM parameters and their functions (Efron & Tibshirani, 1993). Illustrative simulations of stochastic differential equations (Figures 1 and 3D) were based on the Euler-Maruyama method (Higham, 2001).

3.4.3 Analyses for individual symptoms

In addition to sum scores, it would be important to better understand individual depressive symptoms. Regarding statistical estimation, this is a whole different game, however. For example, mBDI involves 21 symptoms, and in order to cover their population distribution with an equally dense sample of observations as for one-dimensional sum score, the amount of required observations scales with exponent 21—which is a lot! For intuition, consider two symptoms assessed with a five-point precision: to get one participant per possible state, one needs at least $5^2 = 25$ participants. To get at least one participant per state spanned by 21 symptoms, one needs $5^{21} > 10^{14}$ participants, and that is more than 10 000 times the entire population of the Earth. This example shows that a brute force empirical exploration of all symptom interactions is unwise (despite symptom correlations reducing the number of feasible states), and hints that understanding symptom interactions is not a task to be exhausted in a single academic thesis; yet, it is an important endeavor in light of sections 1.1 and 1.3. Therefore, three steps were taken to this direction: first, temporal stability of individual symptoms was assessed; second, symptoms associated with chronically rather than transiently high sum score were assessed; and third, causality between sleep and sum scores was assessed from a novel viewpoint.

As we turn to individual items measured with five-point precision instead of their sums, it becomes increasingly important to acknowledge the ordinal measurement scale in most computations. This is perhaps best done using the copula theoretic version of polychoric correlation. Sometimes ordinal variables (*e.g.*, five-point precision

questionnaire items) are observed, even though the underlying phenomenon is likely to involve a pair of continuous variables; polychoric correlation is the correlation between the latent continuous variables as estimated from the ordinal variables. Polychoric correlation was originally introduced by Karl Pearson and based on an assumed underlying Normal distribution (Pearson, 1900). A debate followed about whether or not the assumption of Normality is a shortcoming, but recent evidence shows that the polychoric correlation is not robust for non-Normal underlying distributions; instead, an empirical polychoric correlation that assumes only existence of an arbitrary continuous underlying distribution was proposed (Ekström, 2009). The development draws heavily from the statistical theory of copulas that has evoked recent interest in many areas of application (Genest & Favre, 2007; Nelsen, 2006).

For all bi-variate probability distribution functions H, there exists a copula function C that maps the unit square on plane onto unit interval so that $H(x,y) = C(F(x),G(y))$, where F and G are the marginal distributions for the two random variables X and Y [Sklar's theorem (Genest & Favre, 2007; Nelsen, 2006)]. This means that the essential dependency between any two variables can be modeled independently of their marginal distributions, and is captured by the unique C. If R_i is the rank of observation x_i among n observations of the random variable X, a population version of the normalized rank $R_i/(n+1)$ of the observation $X = x_i$ is given by the "grade" $u_i = F(x_i)$. Spearman's grade, or rank, correlation is then defined, for example (Nelsen, 2006), by

$$\rho = 12 \int_0^1 \int_0^1 C(u,v)\,du\,dv - 3.$$

ρ corresponds to Pearson's product-moment correlation (usually denoted r) of ranks, but ρ obviously is invariant with respect to monotonic transformations of marginal distributions, whereas the standard Pearson's correlation is not (Genest & Favre, 2007). In order to avoid assuming any particular dependency function C, empirical polychoric correlation replaces C with the empirical copula C_n for n observations (Ekström, 2009; Genest & Favre, 2007; Nelsen, 2006). To provide correspondence with the original

polychoric correlation under the special case of Gaussian copula, a minor modification was applied in the form of the map $\rho \rightarrow 2\sin(\rho\pi/6)$ from unit interval to unit interval (Genest & Favre, 2007); and to further correct the bias due to the finite number of five ordinal categories, the output was multiplied with a constant of 1.08, as suggested by Ekström (2009). Hence, when $\rho(C_n)$ denotes the empirical Spearman's grade correlation, the empirical polychoric correlation, r_{epc}, was defined as

$$r_{epc} = 2.16\sin(\rho(C_n)\pi/6). \quad (4)$$

Because the empirical copula C_n is a non-parametric function of data, so is r_{epc} [see Ekström's (2009) thesis for the computational details].

In this thesis, equation 4 was used to compute observed autocorrelations between a depressive symptom measured with a five-point precision at the baseline and the same symptom similarly measured at a later follow-up. If t is the time to the later measurement, the autocorrelation was modeled with a parametric model

$$f_i(t; a_i, K_i) = e^{a_i t} + K_i(1 - e^{a_i t}) \quad (5)$$

where $K_i = \lim_{t\rightarrow\infty} f_i(t;a_i,K_i)$ is the stable level of within-individual similarity in the symptom i, and the negative constant a_i is the relaxation rate to the stable level K_i. Parameters a_i and K_i were estimated from the data, and their 95% confidence intervals were computed by bootstrap methods where needed (Efron & Tibshirani, 1993). As a standard method for such simple estimation tasks, Nelder-Mead simplex method for unconstrained nonlinear optimization was used to minimize the average squared differences (quadratic loss, or 'L^2-loss') between observed empirical polychoric correlations, r_{epc}, for item i and the model estimates, $f_i(t;a_i,K_i)$, with respect to a_i and K_i (the "fminsearch" algorithm in Matlab® software version R2012a, 7.14.0.739 by MathWorks, Inc., Natick, Massachusetts, USA). One may also fit the Equation 5 to the

usual autocorrelations of the mBDI sum score. In addition to studying symptoms' temporal stabilities, their associations with chronicity were assessed.

In order to determine which symptoms associate with chronically rather than transiently elevated dysphoria, or sum score, participants were divided to matched groups as follows. Chronically elevated dysphoria was defined as the condition where the participant belonged to the upper sample tertile of the depression score in all four prospective measurements, that is, in the years 1992, 1997, 2001, and 2008. Transiently dysphoric group was defined by participants exceeding 82.4^{th} percentile cut-off in just a single follow-up; that year's observation was then used in the subsequent comparisons. The cut-off percentile for the transient group was chosen so that the two groups had equal group-average depression sum score. This could have easily been done by hand, but of course, almost any numerical optimization routine performs this simple bounded optimization task with twice the ease and accuracy. A standard solution of Matlab®-software was used again (*i.e.*, the "fminbnd" function based-on golden section search and parabolic interpolation). As a sensitivity analysis, it was verified that both minimizing the absolute difference of group averages ('L^1-loss') and minimizing the squared difference ('L^2-loss') provided the exact same percentile cut-off, 82.4.

With the above definitions, there were 135 chronically dysphoric, 179 transiently dysphoric, and 743 other participants; listwise deletion was applied to missing data, as even flexible imputation methods are not designed to uncover nonlinearities not present in the imputation model (White, Royston, & Wood, 2011), therefore being ill-suited for explorative approaches. The chronic and the transient group had almost the same average and median mBDI score, and very similar age and gender characteristics (see original publication II for comprehensive details). Hypothesis tests of equal symptom averages between chronic and transient groups were computed by permutation test (Efron & Tibshirani, 1993), where the permutation (*i.e.*, comparison) distributions were derived by 10 000 random permutations of the group assignment (chronic, transient, and others). Confidence intervals in figures were computed as 95% Bias-Corrected and Accelerated bootstrap confidence intervals with 2000 bootstrap re-samples (Efron & Tibshirani, 1993). Logistic regression models were applied to estimate predictive

functions for the chronic group membership (Gelman & Hill, 2007). In addition to these uses of longitudinal/panel data, cross-sectional symptom distributions were assessed for causality between sleep problems and other depressive symptoms.

The pairwise causality estimation, as applied here, starts from the assumptions that (1) either sleep problems, x_s, cause depression (the sum of non-sleep related symptoms) or depression, x_d, causes sleep problems, (2) the causal association is linear, (3) independent residual terms are non-Gaussian (distributed according to some other than the Normal distribution), and (4) there are no confounding variables. This is a Linear, Non-Gaussian, Acyclic Model [LiNGAM, (Shimizu et al., 2006)]. Mathematically it means that for centered (zero-mean) variables either

$$\begin{cases} x_s = e_s \\ x_d = b x_s + e_d \end{cases}, \tag{6}$$

or

$$\begin{cases} x_s = b x_d + e_s \\ x_d = e_d \end{cases}, \tag{7}$$

where e_s and/or e_d is a non-Gaussian variable, and b is a constant, non-zero regression coefficient. The aim of the algorithm is to estimate which one holds, the system of equations 6 or the system of equations 7. In these two alternative systems of equations, either depression or sleep problems is an *exogenous* variable: an exogenous variable is not predicted by other variables in the system, and can be considered as an input to a system of variables. The estimated exogenous variable is causal because the other variable is its function, and it is not a function of the other variable. In other words, manipulations of an exogenous variable lead to changes in the other (endogenous) variable, but manipulations of an endogenous variable do not affect the exogenous variable (Pearl, 2000).

With non-Gaussian variables and the LiNGAM model, one may determine causality by estimating which one is the exogenous variable, x_d or x_s, by estimating which one is less dependent on its residuals (Shimizu et al., 2011). In the DirectLiNGAM-algorithm (Shimizu et al., 2011), this dependency is evaluated using a nonparametric kernel-based estimator (Bach & Jordan, 2003) of the mutual information between two variables (Hyvärinen, Karhunen, & Oja, 2001). In addition, other pairwise measures can be constructed (Hyvärinen & Smith, 2013; Hyvärinen, 2010). Three different pairwise measures of causality were applied here: DirectLiNGAM-based, skew-based, and tanh-based (based on hyperbolic-tangent approximation to likelihood ratio for distributions with non-Gaussian kurtosis). For each statistic, a positive value signifies causal antecedence of the first argument/variable, and a negative value indicates the opposite condition. The DirectLiNGAM-based measure applies the default options of the DirectLiNGAM-algorithm version 1.0; that is, the pairwise causality statistic used by the more general DirectLiNGAM-algorithm (Shimizu et al., 2011). The causality algorithms were implemented using the Matlab® software. The three statistics are briefly described below, and although some of them involve intricate calculations, the logic of the skew-based statistic is discussed in some detail in order to provide intuition on how non-Gaussianity may sort out causal directions in the linear model.

First, the DirectLiNGAM-based statistic is described. If one denotes by $M(x_d,x_s)$ the mutual information between x_d and ordinary least squares regression-residual of x_s (estimating e_s in Eq. 7), and by $M(x_s,x_d)$ the mutual information between the opposite configuration, then under the LiNGAM assumptions the inequality $M(x_d,x_s) < M(x_s,x_d)$ implies that x_d is the causal antecedent and *vice versa* (Shimizu et al., 2011). Therefore, one can use the quantity

$$T(x_d,x_s) = M(x_s,x_d) - M(x_d,x_s)$$

as a causality statistic whose positive values indicate that x_d causes x_s, whereas the negative values indicate the opposite causality.The exact same kernel-based pairwise

quantity $M(\cdot,\cdot)$ that the DirectLiNGAM-algorithm uses was applied when deriving causal ordering of variables (Shimizu et al., 2011), referring to this statistic T as the DirectLiNGAM-based statistic; it aims to use general dependency information in variables. More restricted deviations from Gaussianity can also be used for the causality estimation. These include skewness and kurtosis information.

Second, a statistic based on kurtosis of distributions can be directly derived. The simple kurtosis, however, suffers from a lack of robustness and from sign-indeterminacy (Hyvärinen & Smith, 2013; Hyvärinen, 2010). A hyperbolic tangent function (tanh) offers a more useful approximation (Hyvärinen et al., 2001; Hyvärinen & Smith, 2013; Hyvärinen, 2010). The explicit rationale is beyond the present scope, but the ensuing statistic is

$$T_{tanh}(x_d,x_s) = r(x_d,x_s)E[x_d \tanh(x_s) - x_s \tanh(x_d)],$$

where the input variables must be standardized z-scores; r is product-moment correlation of the variables, and E denotes the expectation operator. This is called the tanh-based causality statistic herein.

Third, a simple statistic based on skewness exists. Let variables x_d and x_s be standardized (mean zero, variance one) variables with positive skewness, then the desired skewness-based statistics is

$$T_{skew}(x_d,x_s) = r(x_d,x_s)E[x_d^2 x_s - x_s^2 x_d].$$

The sign-requirement is not a limitation, as if a variable x^* has a negative skewness, then the statistics can nonetheless be applied to $x = \text{sign}(\text{skew}(x^*))x^*$; that is, a skewed variable multiplied by the sign of its skewness always has a positive skew. The statistic can be understood as follows.

If x and y are standardized variables with positive skewnesses and $y = rx + e$ holds, we have $T_{skew}(x,y) = r(E[x^3 r + e] - E[x(rx + e)(rx + e)])$. Using standard calculus for

expectations, independence of x from error e, and the fact that $E[x^3] = \text{skew}(x)$ for a standardized variable x, one easily obtains that $T_{skew}(x,y) = \text{skew}(x)(r^2 - r^3)$. As $\text{skew}(x) > 0$, and $|r| < 1$, it follows that $T_{skew}(x,y) > 0$. But when $x = ry + e$ holds, similar calculations yield $T_{skew}(x,y) = \text{skew}(y)(r^3 - r^2) < 0$. Hence, if x is cause under the linear model, this is detected by the positive values of the statistic $T_{skew}(x,y)$, and the causality from y to x is detected by the negative values.

It can already be seen that T_{skew} is robust against Normally distributed measurement error since it suffices that skewness is non-zero, and also simulations have demonstrated the robustness against measurement errors for the three statistics (Hyvärinen & Smith, 2013). The original contribution III further demonstrated reasonable tolerance for partial confounders, but owing to space limitations these results are omitted from the thesis text. So far, little research on missing-data methods exist for these novel causality methods, and hence only complete observations were used (Table 2 for details).

3.4.4 Replicator dynamics, Moran process, and slow-time Markov chains

When trusted by significant others, most people feel bad to let the others down. But where does this 'good nature' come from, if natural selection operates on individual fitnesses enhanced by selfishness? The derivation of parsimonious answers to this question belongs to the joint research domain of game theory and population dynamics, where the population dynamics can be due to genetic evolution and/or due to social learning (Nowak, 2006, 2012; Sigmund, 2010). Regarding the mathematical developments in that research field, imitation of more successful behavioral strategies tends to obey the same or similar formulas as the reproduction of more successful genes, *mutatis mutandis* (Nowak, 2006; Sigmund, 2010).

Assume that a finite number of behavioral staregies, $e_1, e_2, \ldots, e_n$, are available, and proportions $x_1, x_2, \ldots, x_n$ of the population are using them. Assume further that, from time to time, a randomly chosen individual randomly samples (or observes) a strategy from the population, and imitates that strategy with a certain likelihood ("randomly" is

based on current strategy-use frequencies). Then, if for an individual using e_j, the probability of switching from strategy e_j to e_i within a time-frame Δt is given by $x_i f_{ij} \Delta t$, the net influx-outflux to strategy i on interval Δt is described by

$$x_i(t + \Delta t) - x_i(t) = \sum_j f_{ij}\, x_i(t) x_j(t) \Delta t - \sum_j f_{ji}\, x_i(t)\, x_j(t) \Delta t,$$

where $\sum_j$ refers to a sum over index j running from 1 to n. Taking the limit $\Delta t \to 0$, yields the differential equation

$$dx_i/dt = x_i \sum_j (f_{ij} - f_{ji}) x_j. \qquad (8)$$

In order to model population dynamics, the rates f_{ij} generally depend on the state of the population, that is, on the values of $x_1, x_2, \ldots, x_n$. To this end, the "pay-off" matrix from game theory is introduced (Sigmund, 2010).

The pay-off structure of a public good game can be captured to a matrix $A = (a_{ij})_{i,j}$, so that the average income of a participant or player i in the population state $(x_1, x_2, \ldots, x_n) = x$ is $\sum_j a_{ij} x_j = (Ax)_i$, where the latter term denotes the i^{th} element of the vector resulting from the product of the pay-off matrix A and the column-vector of states x. It is then customary to assume that $f_{ij} = \max\{(Ax)_i - (Ax)_j, 0\}$; in other words, a player may adopt another strategy only if it provides a higher pay-off than the current strategy, and the possible switch gets more probable as the difference in pay-off does. In this case, $(f_{ij} - f_{ji}) = (Ax)_i - (Ax)_j$, and the Equation 8 becomes

$$dx_i/dt = x_i \sum_j [(Ax)_i - (Ax)_j] x_j = x_i [(Ax)_i - x \cdot Ax], \qquad (9)$$

where $x \cdot Ax$ is a dot/inner product of the two vectors; $x \cdot Ax$ also represents the average income of the joint enterprise at the given time. Equation 9 is known as the *replicator equation*, and it can be interpreted in two ways (Sigmund, 2010): (1) the per capita growth rate is a function of a strategy's pay-off relative to the average of all strategies, or (2) the change in strategy i is proportional to the product of relative pay-off of the

strategy and prevalence of that strategy. The latter could be further interpreted so that the proportion of a strategy in population depends not only on its relative success but also on the likelihood of observing it, observations thereby constraining the chance of imitating or learning a new strategy.

Replicator equation involves theoretically convenient infinite populations: only proportions of strategies are modeled instead of the absolute numbers of players. Studying finite populations offers a complementary approach (Nowak, 2006; Sigmund, 2010). For finite populations, random fluctuations, due to sampling effects, for example, need to be taken into account; instead of deterministic differential equations, stochastic models are implied. Moran process is a natural finite-population analogy of the replicator equation (Nowak, 2006). It is assumed that i individuals of type/strategy e_1 and $M - i$ type e_2 individuals exist in a population of M individuals (i now indexes states instead of the two strategies). Moran process is a discrete-time random process: in each time step, single individual is chosen for death (or abandons a strategy for another one) and single individual is chosen to be replicated or imitated. The transition matrix, yielding probabilities $p_{i,j}$ of moving from population-state with i type e_1 individuals to a state with j type e_1 individuals, is tri-diagonal ($p_{i,j} = 0$, if $|i - j| > 1$). Influx to the type e_1, or 'birth rate' for e_1, is

$$p_{i,\,i+1} = \left(\frac{M-i}{M}\right)\left(\frac{if_i}{if_i + (M-i)\,g_i}\right),$$

where f_i is the fitness of type/strategy e_1 in the population-state i, and g_i the fitness of e_2. The first fraction is the probability that the updating (or 'dying') individual is of type e_2, and the second fraction is the probability that the imitated/selected type is e_1. Outflux from type e_1 (influx to e_2) is then

$$p_{i,\,i-1} = \left(\frac{i}{M}\right)\left(\frac{(M-i)\,g_i}{if_i + (M-i)\,g_i}\right).$$

The population-states 0 and M are 'absorbing' states, where only single strategy is present and therefore cannot be selected against. Assuming that

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

is the pay-off matrix of a two-strategy game, then the expected pay-off for the strategy e_1 players in the population-state i is

$$F_i = \alpha\left(\frac{i-1}{M-1}\right) + \beta\left(\frac{M-i}{M-1}\right),$$

whereas that of for the e_2 player is

$$G_i = \gamma\left(\frac{i}{M-1}\right) + \delta\left(\frac{M-i-1}{M-1}\right).$$

We then assume that the parameter $s \in [0,1]$ measures the importance of the game for overall success, and thereby the strength of selection. Then the above fitnesses become $f_i = 1 - s + sF_i$ and $g_i = 1 - s + sG_i$. In sufficient time, this stochastic process will inevitably drift into one or other of the absorbing single-strategy states (Klenke, 2008). The probability of fixation, or 'absorption', to the state with all M players using the first strategy, after introducing just a single player of first strategy (an innovation) into the population of second-strategy players, is

$$\rho_{1,2} = 1 / \left(1 + \sum_{k=1}^{M-1} \prod_{i=1}^{k} \frac{1-s+sG_i}{1-s+sF_i}\right). \tag{10}$$

That is, $\rho_{1,2}$ is the probability that a single player of strategy e_1 who is introduced to a population full of strategy e_2 players is eventually copied by the entire population (Sigmund, 2010).

In addition to imitation, it is customary to assume a certain rate of innovations, or behavioral exploration, where a novel strategy is spontaneously introduced to the population; in the genetic setting, both the replicator equation and the Moran process would model selection, and the innovations would correspond to mutations (Nowak, 2006; Sigmund, 2010). Importantly, when the rate of innovations (respectively, mutations) is very low compared to the temporal dynamics of social imitation (resp. selection), cultural (resp. genetic) evolution in the finite populations can be approximated with a discrete-time Markov Chain [a stochastic process that does not depend on its past given the present (Klenke, 2008; Nowak, 2006; Sigmund, 2010)].

Dynamics under rare innovations can be approximated by assuming that after each innovation, the finite population has had the "sufficient time" to randomly drift into an absorbing state where only one strategy is present, before the next innovation/mutation introduces a single individual with another strategy; initiating a new competitive process between the prevalent strategy and the newly innovated one. This is a slow-fast time-scale separation, and justified when imitation/selection operates much faster than innovation/mutation. The probabilities of the end result of competition can be computed by this approximation from the Equation 10, and from these, the time-average prevalence of the strategies.

A Markov chain on the dominant (slow-time scale) population states is commonly modeled using fixation probabilities (Eq.10) of the general Moran process (Nowak, 2006; Sigmund, 2010). The Moran process unfolds in the *fast-time scale*. The approximating Markov chain progresses in the *slow-time scale* of discrete single-strategy states, whose transition probabilities are given by Equation 10. The Equation 10 provides the link between the imitation dynamics and time-averages of population states, through the general theory of Markov chains (Klenke, 2008).

When genetic selection is modeled instead of the social imitation, "chosen for imitation" is replaced with "chosen for reproduction", and "first strategy" can be replaced with "presence of the genotype in question", whereas "second strategy" refers to "lack of the genotype". Hence, $\rho_{1,2}$ is the transition probability from *'no-gene'* state to a state with the *'gene'*, $1 - \rho_{1,2}$ is the probability of staying in the *no-gene* state (mutation introduces the gene, but it goes extinct before taking over the population), $\rho_{2,1}$ is transition probability from *gene* state to *no-gene* state (a mutation deletes the gene from population; Eq. 10 with positions of F_i and G_i reversed), and $1 - \rho_{2,1}$ is the probability of staying in *gene* state (deletion does not take over). The eigenvector of the standard Markov-Chain theory transition matrix that is associated with the maximal eigenvalue (*i.e.*, one) yields the average times that the chain spends in each state, or probabilities of finding the chain at a given state in a randomly chosen time: P(*gene*) = $\rho_{1,2}/(\rho_{1,2}+\rho_{2,1})$, and P(*no gene*) = $\rho_{2,1}/(\rho_{1,2}+\rho_{2,1})$. In such a model, the time taken by the selection process is considered negligible compared to mutation occurrences, and the

time-scale of primary interest is counted on discrete mutations (which is the "slow time"). Similar well-understood theory holds for the case with more than two available states (Klenke, 2008; Sigmund, 2010).

When several strategies of interest exist, several Moran processes need to be defined, one for each pairing of a dominant strategy and an invading strategy. If the public good game of the section 3.3.1 is modeled using the finite-population Moran processes, one replaces fitnesses F_i and G_i with, for example, average pay-off of Punishers who are playing against Defectors in the population of i Punishers and $j = M - i$ Defectors, denoted by F_i. The pay-off for Defectors is then G_i. In a population of size M with i Punishers and $M - i$ Defectors, the probability to select k Punishers and $N - k$ Defectors in N trials (recall that N players enter to joint enterprise of the section 3.3.1) is given by the hypergeometric distribution

$$H(k,N-1,i,M) = \frac{\binom{i}{k}\binom{M-i}{N-k}}{\binom{M}{N}}.$$

Then, in the population of i Punishers and $j = M - i$ Defectors, the average pay-offs F_i and G_i are given by

$$F_i = \sum_{k=0}^{N-1} H(k,N-1,i-1,M-1)\left(\frac{k+1}{N}rc-c\right) = \frac{rc}{N}\left(1+(i-1)\frac{N-1}{M-1}\right)-c, \qquad (11)$$

and

$$G_i = \sum_{k=0}^{N-1} H(k,N-1,i,M-1)\left(\frac{k}{N}rc\right) = \frac{rc(N-1)}{N(M-1)}i. \qquad (12)$$

These and similar tedious, but straightforward, calculations for all pairings of all available strategies have been provided in previous research, as well as the transition matrices that result from applying Equation 10 to them (Hauert et al., 2007; Sigmund, 2010).

According to the general theory of Markov chains (Hauert et al., 2007; Klenke, 2008; Sigmund, 2010), the transition matrix

$$\begin{pmatrix} 1-\rho_{2,1}-\rho_{3,1}-\rho_{4,1} & \rho_{1,2} & \rho_{1,3} & \rho_{14} \\ \rho_{2,1} & 1-\rho_{1,2}-\rho_{3,2}-\rho_{4,2} & \rho_{2,3} & \rho_{2,4} \\ \rho_{3,1} & \rho_{3,2} & 1-\rho_{1,3}-\rho_{2,3}-\rho_{4,3} & \rho_{3,4} \\ \rho_{4,1} & \rho_{4,2} & \rho_{4,3} & 1-\rho_{1,4}-\rho_{2,4}-\rho_{3,4} \end{pmatrix}$$

constructed by applying Equation 10 to all pairwise pairings of strategies' pay-offs admits a stationary distribution describing the average time spent in each strategy/state (*e.g.*, $\rho_{i,j}$ signifies absorption/fixation probability of entire population to strategy i after introducing single player of that strategy into a population full of j players, where i and j run through set of strategies S = {Cooperator, Defector, Nonparticipant, Punisher}). This means that after sufficiently many innovations, if an observation is made at a random time, the probability of finding the population in the Cooperator state is precisely known, and given by the first element of maximal eigenvector for the above transition matrix. Similarly, probabilities of all states are known (assuming the negligible "fast" time-scale), and correspond to the temporal average of ("slow") time spent in those states, as given by ergodic theorems (Klenke, 2008). In fact, Markov chains represent one of the few cases where instantaneous probabilities (*c.f.*, between-individual distributions for independent individuals) and temporal averages (*c.f.*, within-individual processes) correspond, although similar equivalences are frequently assumed even when they are not justified (Molenaar & Campbell, 2009; Molenaar, 2004).

Let the stationary distribution (relative time-averages) of cooperative, defective, nonparticipating, and punishing strategies in the public good game of section 3.3.1 be denoted by (π_c, π_d, π_n, π_p), and let (π'_c, π'_d, π'_p) be the stationary distribution of the

otherwise similar game that lacks the nonparticipating strategy. The stationary distributions (π_c, π_d, π_n, π_p) and (π'_c, π'_d, π'_p) are derived exactly as in previous reports (Hauert et al., 2007; Sigmund, 2010), and re-stating the formulas is therefore avoided. The Results section examines these stationary distributions, and the nonparticipating strategy is interpreted as sufficiently elevated depressive symptoms. Because depression is seen as a part of a biologically ingrained behavioral repertoire, usually being an involuntary reaction rather than a conscious choice, it becomes important to demonstrate that also genetic evolution favors such strategy.

The stationary distribution (π_c, π_d, π_n, π_p) models states of a closed system of four strategies, but the scientific premise of parsimony requires a justification for each and every added layer of complexity in a model. Taking cooperative, defective, and punishing strategies as given, this study aims to demonstrate that evolution will favor the emergence of the fourth behavioral strategy of nonparticipation; behavioral strategy refers here to a robust and general mode of behavior that has a biological basis, not only to an abstract conscious strategy. Conscious cortical processing does not provide very efficient means for coping with important biosocial threats, as can be seen, for example, in the frequent injuries suffered by those who cannot sense pain or discomfort (Nesse, 1991). Furthermore, evolution must build on existing 'facilities'. Indeed, it is often difficult to separate emotional and physical pain, and it has been argued that this is because the physical pain system was already in place when social animals evolved adaptations for responding to social threats (MacDonald & Leary, 2005). Still furthermore, intergroup competition and conflicts appear empirically as a potent evolutionary phenomena (Bowles, 2009; Puurtinen & Mappes, 2009; Sääksvuori, Mappes, & Puurtinen, 2011), and therefore ought to be incorporated into comprehensive models.

According to the present model, gains from joint enterprises drive a natural selection process between depression-vulnerable, or -capable, genetic variant (Nonparticipation strategy) existing in population *versus* not existing; the idea is that such a genetic variant is introduced to populations on rare occasions (with a low mutation rate) and affects to the competition against other populations, and therefore the survival of the

population. Individuals in each group/population occasionally engage in a joint enterprise, and the time-average of the group's average income from those enterprises affects the group's genetic fitness proportionally to the genetic selection strength, s_g. The modeled 'genetic variant' refers to capacity for the fourth nonparticipation strategy, whose final expression in the population depends on the 'environment' of other strategies. Therefore, the modeled totality represents an interaction of genes and environment.

Let us model a relative phenomenon, standardizing by $c = 1$; then, the time-averaged outcome from the joint enterprise of section 3.3.1 is:

$$F^* = \pi_c(r\text{-}1) + \pi_d 0 + \pi_n\sigma + \pi_p(r\text{-}1) = (\pi_c+\pi_p)(r\text{-}1) + \pi_n\sigma.$$

When nonparticipating strategy is not available and (π'_c, π'_d, π'_p) is the stationary distribution of the three available strategies, the average pay-off is:

$$G^* = (\pi'_c+\pi'_p)(r\text{-}1).$$

We normalize as, $F = F^*/(F^* + G^*)$ and $G = G^*/(F^* + G^*)$, so that s_g can be directly interpreted as the relative selection strength of environment for the depression genotype (*i.e.*, for availability of a nonparticipation strategy). The success of strategies compared to the other individuals' strategies *within* a group affects the intensity of social imitation proportionally to imitation (strategy-selection) strength denoted by s_i [corresponds to the selection strength parameter s in previous studies of the public good game (Hauert et al., 2007; Sigmund, 2010)]. Hence, the pairwise pay-offs (*e.g.*, Eq. 11 and 12) and s_i are substituted for F_i, G_i, and s in Equation 10 when modeling imitation dynamics, whereas the F, G, and s_g replace them when modeling slower progressing genetic evolution.

When all the four strategies are available, the joint enterprise of section 3.3.1 is not readily amenable for infinite-population analysis due to unstable solutions; but for three strategies, an infinite-population model based on replicator dynamics (Eq. 9) yields inferences that are consistent with the approximate model using the Moran process (Eq.

10) and a Markov chain of ("slow") innovations (Sigmund, 2010). Simulation studies indicate that the approximation is reasonably accurate for the four strategy situation as well (Hauert et al., 2007). Sigmund's (2010) equations for three-strategy replicator dynamics are visually examined in the Results section, as they provide intuition on the 'rock-paper-scissor' dynamics found for the three strategies. Primary interests in this study, however, are the probability with which a nonparticipating strategy is found from the wider population (among all groups) compared to the state without such evolved mode of behavior (i.e., the quantity $\rho_{1,2}/(\rho_{1,2}+\rho_{2,1})$ and the existing empirical evidence connecting nonparticipation and depression. All related computations and figure drawing were performed using the Matlab® software.

4 RESULTS

Figure 2 shows the distributions of standardized (z-score) depression inventories (Figure 2A), the quadratic relationship between mBDI and BDI-II in Young Finns data (Figure 2B), and the relative information-value per severity for the both scales, together with normalized factor-score distribution that shows what severity states mainly existed in the data (Figure 2C). The lower row of panels (Figures 2D-F) shows the relationships between depression inventories and sleep-problem scores; here, two sleep-related items regarding changed amount of sleep (BDI-II) and subjective feelings of tiredness (mBDI and BDI-II), were removed from the depressive-symptom sums/averages prior to the z-score transformation. The same removal was in effect in Table 2.

It often happens in population studies that some part of the data is available for a given participant and another part is not. Comparing those with all data against those with partial data is a simple method to gain some idea about the dominant direction of attrition-related sampling bias, even though nothing is known about those who lack all the data. Table 2 both provides some basic descriptives of the evaluated data sets and assesses the attrition bias via the above-mentioned simple method. The study group and the attrition group were formed with respect to having *versus* not having either sleep-

problems score or depressive-symptoms score. Only the 2008 follow up with both mBDI and BDI-II is shown from the Young Finns study. The higher amount of attrition in the Young Finns study compared to the Wisconsin study may be due to the many follow ups of the Young Finns study that could have been taxing for some participants; as well as the separate questionnaires for depression and sleep problems due to the different coordinating study centers.

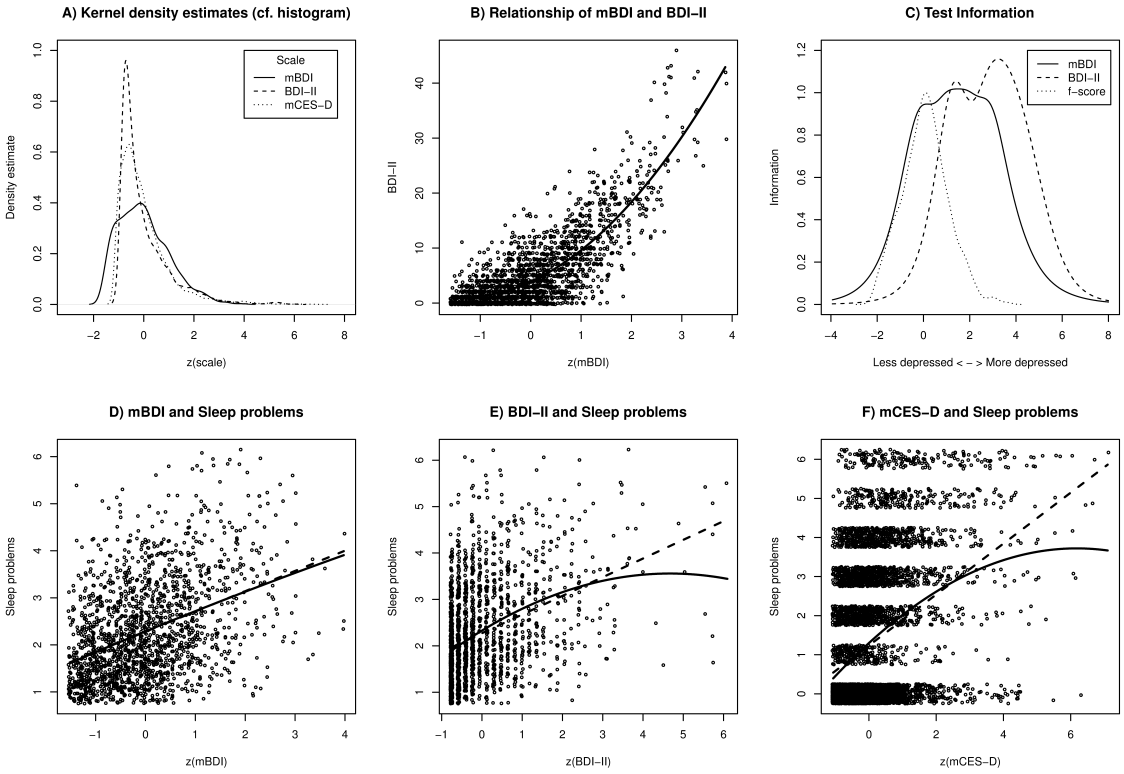

Figure 2. Depression Inventory Distributions. (A) Kernel density estimates using a Gaussian kernel.(B) Scatterplot of BDI-II values against z-score of mBDI (points) and an ordinary least squares regression estimate with a quadratic term (line). (C) Estimated test/scale (Fisher) informations under Graded Response Model applied simultaneously to items of both scales: "f-score" refers to re-scaled density (maximum of 1) of participants' factor scores under the model, showing where the bulk of severity levels in the data lies (*y*-axis values are relative and hold no absolute meaning outside the figure). (D-F) Scatterplot (points), linear model (dashed line), and linear model with quadratic term (solid line), for the sleep-problem and depression scores. In all scatterplots, a uniform random value (a.k.a. 'jitter') between -1/40 and 1/40 was added to both *x*- and *y*-axis values (circle locations) in order to improve visibility of the data points.

Table 2. Some Basic Sample Characteristics

Comparisons between mBDI (1. depression scale) and sleep problems in Young Finns data						
Measure	**Study sample**		**Attrition sample**		***p*-value**	***d*-value**
Number of participants	1699		1897			
Percentage of males	41.1%		56.2%		< 0.001	
	mean	range	mean	range		
Age of participants (years)	38.71	31-46	38.20 (n = 1897)	31-46	0.002	
	mean	s.d.	mean	s.d.		
Sleep problems score (1-6)	2.28	1.05	2.31 (n = 463)	1.06	0.543	0.029
Depression score (1-5)	2.00	0.66	2.14 (n = 333)	0.65	< 0.001	0.213
Comparisons between BDI-II (2. depression scale) and sleep problems in Young Finns data						
Measure	**Study sample**		**Attrition sample**		***p*-value**	***d*-value**
Number of participants	1687		1909			
Percentage of males	40.9%		56.2%		< 0.001	
	mean	range	mean	range		
Age of participants (years)	38.67	31-46	38.24 (n=1909)	31-46	0.011	
	mean	s.d.	mean	s.d.		
Sleep problems score (1-6)	2.27	1.04	2.34 (n = 475)	1.07	0.186	0.067
Depression score (0-3)	0.23	0.30	0.54 (n = 328)	0.64	< 0.001	0.637
Comparisons between mCES-D (3. depression scale) and sleep problems in Wisconsin data						
Measure	**Study sample**		**Attrition sample**		***p*-value**	***d*-value**
Number of participants	6640		3677			
Percentage of males	46.7%		51.4%		< 0.001	
	mean	range	mean	range		
Age of participants (years)	53.14	52-55	53.19 (n = 3084)	52-55	< 0.001	
	mean	s.d.	mean	s.d.		
Sleep problems score (1-6)	1.24	1.75	0.63 (n = 90)	1.48	0.001	0.349
Depression score (0-140)	16.40	15.44	23.31 (n = 167)	19.81	< 0.001	0.219

Note: *p*-value is from *t*- or chi-squared test for the difference between the study and attrition samples, *d*-value is the Cohen's effect size of the mean difference, and "s.d". denotes standard deviation. Attrition sample consists of participants who lacked information either regarding depression or regarding sleep. Some had one but not other, allowing comparison against those with both. For such cases, n denotes sample size for this sub-sample. Abbreviations are "mBDI = modified Beck's Depression Inventory", "BDI-II = Beck's Depression Inventory II", and "mCES-D = modified Center for Epidemiologic Studies Depression scale".

Several relevant observations follows from the Figure 2 and Table 2. Figure 2A clearly shows that the distribution of mBDI is less skewed and less kurtotic than the distribution of mCES-D, but the mCES-D variable is still less skewed and kurtotic than the BDI-II. Furthermore, the relationship between mBDI and BDI-II was not linear, but the BDI-II de-emphasizes the mBDI values below plus one standard deviation from mean and expands those above (see the quadratic relationship in Fig. 2B). When treated linearly, as in standard Item Response Theory, BDI-II appears to encode more severe depressive symptoms than mBDI; however, most participants in the data set did not reach the severity levels optimally encoded by the BDI-II (Fig. 2C). Because BDI-II emphasizes severe symptoms, and because high attrition was associated with high levels of depressive symptoms, attrition had thrice the effect on average levels of BDI-II compared to mBDI (Table 2). These properties may also be reflected on the associations between depression scores and sleep problems, as the association was much more linear for mBDI than for BDI-II and mCES-D (Figures 2D-F), and the nonlinearities observed for BDI-II and mCES-D were to a corrective direction compared to that between mBDI and BDI-II (*cf.*, Fig. 2B).

Together the above findings suggest that, among the three inventories, mBDI most comprehensively captures the variance relevant for the indicated general populations. This view is further supported by the fact that almost all reliable variance in BDI-II can be predicted from mBDI. Reliability of sum scores, that is, the correlation between 'true' score and the sum scores, is most commonly estimated with Cronbach's alpha (Lord & Novick, 1968), which was 0.88 for the BDI-II. Yet, the quadratic estimate based on mBDI correlated with BDI-II at 0.83 [*i.e.*, the square-root of adjusted (classical) coefficient of determination was 0.83 (Gelman & Hill, 2007)]. This is a fortunate state of matters, as mBDI has been recorded four times during the Young Finns study, whereas BDI-II only once. In general, this thesis emphasizes mBDI over the other measures, but uses all the inventories when possible.

4.1 Time trajectories of a sum score

Depressive symptom sum, as given by the mBDI, was reasonably close to normal distribution (Fig. 2A). Indeed, sample kurtoses of mBDI in the four Young-Finns follow-ups varied between 2.89 and 3.24 (Normal distribution is 3), and sample skewnesses were all about -0.62 (Normality is 0). Although in an absolute sense all distributions significantly differed from Normal distribution ($p < 0.001$ for Lilliefors test), this small deviations from Normality are not generally problematic for FIML estimation, and FIML estimation tends to outperform simpler standard approaches for missing-data handling in the presence of non-Normality (Enders, 2001). Hence, estimation of EDM via the FIML method was a feasible option. Although there is no natural scaling for sum scores, mBDI (range 21 to 105) was translated and re-scaled to vary between 0 and 63 similar to the more widely used BDI-II; of course, the distributions of scales still differed despite the affine transformation. Prior to the EDM estimation, potential differences in covariance structure among the age cohorts and between men *versus* women were also tested for.

In order to test whether all six Young Finns cohorts could be treated together or not, a likelihood-ratio test was performed for a model with equal means and covariances for all cohorts versus the multi-group 'saturated/unrestricted model' with means and covariances varying freely across cohorts. This constitutes a test for equality of mean and covariance structures between cohorts, with appropriate handling of the missing data achieved via the full information estimation. The difference between cohorts was nearly significant ($\chi^2 = 89.395$, $df_{difference} = 70$, $p = 0.059$), but BIC and AIC supported treating them as a one set of data rather than as separate cohorts (-19051.17 *vs.* -18567.43 and 32850.23 *vs.* 32900.84). Minor differences among cohort means might inflate the trait component of EDM, and thus the same test was done after translating all cohort means to the grand mean; besides, one does not need EDM to interpret simple cohort mean differences. This retained the time variation but removed overall mean

differences among cohorts. After such adjustment it was safer to conclude that all cohorts can be analyzed together ($\chi^2 = 85.863$, $df_{difference} = 70$, $p = 0.096$). Again, BIC (-19010.18 *vs*. -18522.92) and AIC (32891.22 *vs*. 32945.35) supported this choice.

After the removal of cohort means, another test assessed whether it was possible to fit a single model combining men and women. The data suggested that solutions would be different for men and women, regardless of excluding the overall mean levels of sexes ($p < 1.362 \cdot 10^{-11}$ with or without). Thus, all further results will be given separately for men and women. This solution was also favored by AIC (43299.11 *vs.* 43352.82), and previous literature on depression (Hankin et al., 1998; Hyde et al., 2008); but, not by BIC (-8515.67 *vs.* -8548.58) that penalizes more heavily from excess parameters (Kass & Raftery, 1995). The Table 3 provides the saturated FIML covariance estimates. Although the full-information estimation yielded higher average mBDI scores than typical listwise deletion (Enders, 2001) of missing observations, the effect was vanishingly small, without systematic trends by year; on average, Cohen's *d* effect size was 0.021 for men and 0.015 for women. This was despite the fact that only 177/1832 (9.6%) women lacked all four follow-ups of mBDI compared to 1133/1832 (61.8%) women who lacked some years observation, and 358/1764 (20%) of men lacked all observations compared to 1406/1764 (79.7%) with a partial data. Altogether, the FIML approach could use data from 3061 participants, whereas the listwise deletion would have resulted in only 1057 participants. Marked differences were lacking for covariance estimates as well. Despite these findings, FIML was nonetheless applied where possible, as it is theoretically superior to listwise deletion (Enders, 2001; Muthén et al., 1987).

Table 3. Full Information Maximum Likelihood Estimates of Depression Score (mBDI) Means and Covariance in the Young Finns Data

	Women (n=1832)				**Men** (n=1764)			
Years	**1992**	**1997**	**2001**	**2008**	**1992**	**1997**	**2001**	**2008**
$\mathbf{n}_{missing}$	592	594	635	633	758	940	899	931
Means	19.48	19.15	18.10	17.66	16.63	16.51	15.34	16.09
Covariance								
1992	91.55	-	-	-	84.31	-	-	-
1997	57.18	111.84	-	-	62.28	105.62	-	-
2001	56.32	74.92	124.79	-	49.77	69.00	98.05	-
2008	53.24	60.64	77.87	120.85	49.41	64.53	71.70	105.08

Note: Values on the diagonal are variances, other values are covariances among the measurements.

Based on the above comparisons, EDM was estimated separately for men and women. According to BIC and AIC, EDM superseded the unrestricted model for women (EDM_{BIC} = -11754.06 *vs.* $unrestricted_{BIC}$ = -1323.357; EDM_{AIC} = 25173.11 *vs.* $unrestricted_{AIC}$ = 25470.62). The situation was similar for men (EDM_{BIC} = -11379.91 *vs.* $unrestricted_{BIC}$ = -1421.58; EDM_{AIC} = 17551.79 *vs.* $unrestricted_{AIC}$ = 17818.76). Thus, from the model parsimony point of view, EDM appears to fit to the data in the absolute sense; evidence for EDM compared to the unrestricted model was very strong (Kass & Raftery, 1995). The exact EDM parameter estimates are available in the original contribution I, and here Figure 3 summarizes the main qualitative findings from exact discrete modeling.

Figure 3A illustrates the estimated relative proportions of variance due to stable trait, cumulative sources, and state/error; it shows the proportional long run (infinite-time) asymptotes of the respective components. Limiting values are more informative about the underlying population values, because the data contains no information about how participants have arrived to the depression value observed at the initial measurement; however, Figure 3B demonstrates that these limiting values are approached in a

reasonable time-frame compared to human life-spans. For women, 32.2% of the variance (90% confidence-interval = (18.5%, 54.2%)) was explained by stable trait-like component, 43.9% (23.6%, 58.9%) by sources that accumulate in time, and 23.9% (19.1%, 27.3%) by state-fluctuation or measurement error. For men, 61.4% (48.9%, 69.2%) of variance was due to stable trait, 23.0% (9.9%, 41.7%) due to cumulative sources, and 15.6% (5.6%, 23.2%) due to state-fluctuations. Precise decomposition between trait and cumulative sources was more difficult to obtain than the state variance, as can be seen from the crude confidence intervals.

In addition to the variance decomposition, Figure 3C shows the population-mean trajectories implied by the EDM estimates. Clearly, the trajectories were declining by age, and women had more depressive symptoms than men on average; in addition, women had more variance in mBDI than men (Table 3). The Cohen's effect size of gender on mBDI was small, and generally declined as the participants aged (d = 0.30 in 1992, d = 0.25 in 1997, d = 0.26 in 2001, and d = 0.15 in 2007; see Fig. 3C). Finally, it is worth observing that the estimated model is realistic in the sense that it incorporates the mean-reverting property, or hedonic adaptation ("treadmill"), often observed in real data. The non-state part of the EDM, $x(t)$, corresponds to a stochastic differential equation similar to Ornstein-Uhlenbeck process (Oravecz, Tuerlinckx, & Vandekerckhove, 2011; Øksendal, 2003); its trajectories tend to drift back towards the individual's natural base-level of content. This is illustrated in Figure 3D, where the upper panel draws ten simulated paths of estimated process, $x(t)$, and the lower panel draws ten paths from its driving random process, $GW(t)$, without the additional feedback structure of the model. Future drift of a Brownian motion (lower panel) is independent of its current state, and therefore the plain Wiener process is an insufficient model for the returning to base-level of affect after a perturbing life-event; although some paths might behave realistically, the total model would seem unrealistic for depressive symptoms in the general population compared to the full EDM (Figure 3D, upper panel).

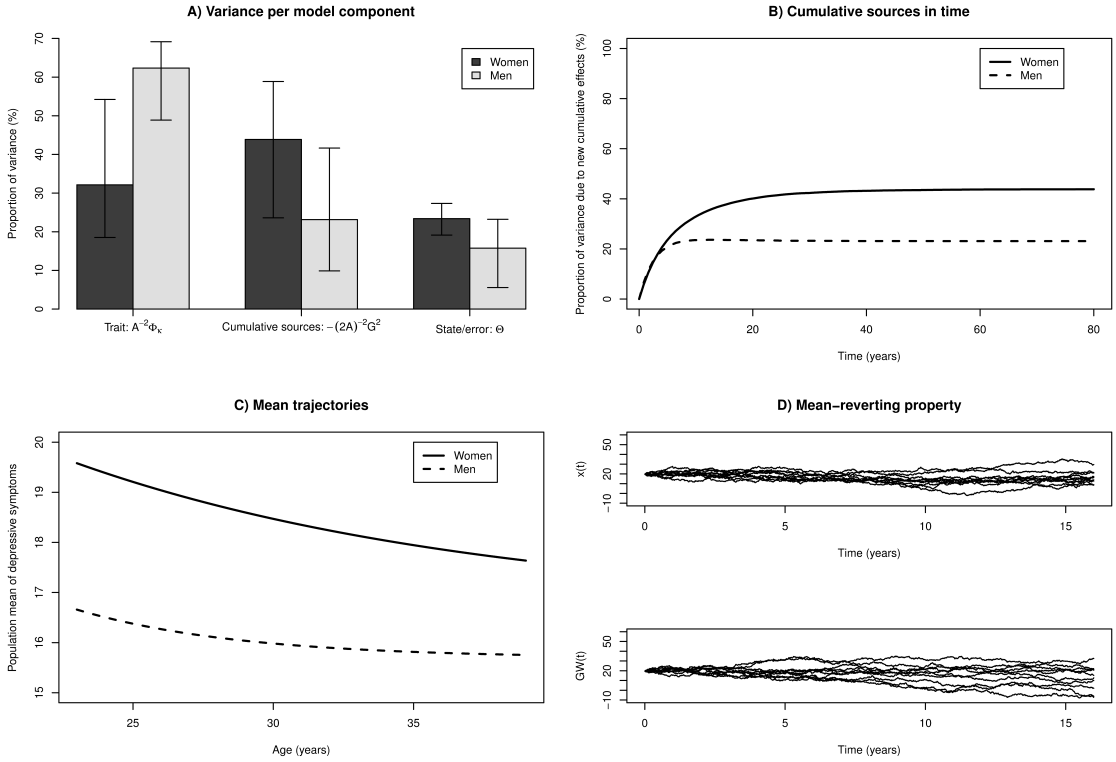

Figure 3. Results from Exact Discrete Model (EDM) estimate. (A) Variance due to trait-like, time-accumulating, and state-like sources at the infinite-time limit. (B) Proportion of variance expected due to novel cumulative sources (after a time zero) as a function of passing time. (C) Model-implied population-mean trajectories. (D) Illustration of the mean-reverting property of the estimated EDM: upper panel shows ten random paths of estimated process *x(t)* (state variance not included), and lower panel shows ten random paths of the Wiener process *GW(t)*; the former having more realistic spread in their drift than the latter.

More specifically, the Ornstein-Uhlenbeck process is typically expressed by the stochastic differential equation

$$dx(t) = \beta(\mu - x(t))dt + GdW(t),$$

where β *and* G are positive constants, and μ describes the baseline level of affect towards which random perturbations eventually tend to with a speed proportional to β (Oravecz et al., 2011; Øksendal, 2003). The parameter A in the EDM model (Eq. 3) was estimated to have a negative value ($A = –0.068$ for men, and $A = –0.172$ for women). Therefore we may write for a single participant i that $\beta = –A$ and $\mu = (b + \kappa_i)/\beta$, where β is then a positive constant and κ_i a fixed realization of a trait value, yielding constant μ. Hence, the continuous, non-state, process $x_i(t)$ of the EDM (Eq. 3) is, for a single person, precisely the Ornstein-Uhlenbeck process.

4.2 Symptom-level results

Having analyzed distributions and temporal trajectories of depressive-symptom sum scores, this thesis proceeds to analyze the trajectories of the individual depressive symptoms. Figure 4 shows the polychoric autocorrelations of depressive symptoms (circles) and the simple model (Eq. 5) estimates for the autocorrelation functions (lines); clearly, the applied model explains the major part of variance in the autocorrelations. Observing the figure, one sees, for example, that autocorrelation of "body-image dissatisfaction" stayed relatively high as the time passed compared to the other symptoms. In contrast, symptoms such as "morning tiredness" and "irritability" had low autocorrelations, and therefore low levels of temporal stability. The autocorrelation of body-image dissatisfaction was estimated to asymptote at $K_{14} = 0.39$, whereas that of the irritability asymptoted at $K_{11} = 0.23$. Also, autocorrelation of the sum score is shown for comparison ($K = 0.51$).

Figure 5A shows the 95% bootstrap confidence intervals for the differences between the estimated long-term stability (model parameter K_i in Equation 5) of body-image dissatisfaction *versus* long-term stability of the other symptoms. The differences are ordered by magnitude in Figure 5A, and one can observe that body-image

dissatisfaction was statistically significantly more stable than most other symptoms; although few other symptoms displayed a comparable stability.

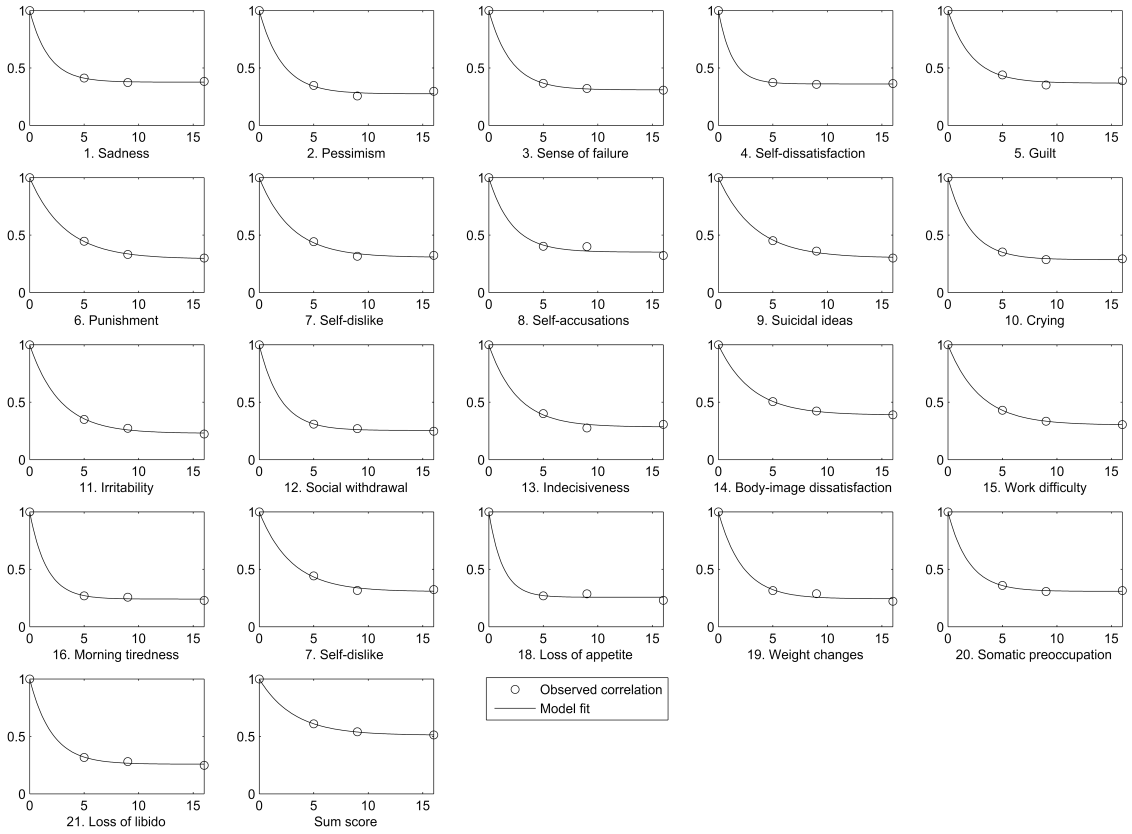

Figure 4. Polychoric autocorrelations of depressive symptoms in the mBDI inventory. Horizontal axes represent time in years, circles denote empirically observed correlations with the measurement at the time zero (year 1992 follow-up of Young Finns study), and lines represent estimated simple two-parameter models (Eq. 5) for the autocorrelations of symptoms. The last panel shows autocorrelations for the mBDI sum score.

To test whether some symptoms associated more strongly with a chronic rather than transient dysphoria compared to the other symptoms, two groups of individuals were matched in terms of their average levels of depressive symptoms: those with chronic and those with transiently high scores over the four follow-ups that spanned 16 years

altogether (section 3.4.3 for exact group definitions). Figure 5B shows the difference between the average item values in chronic and transient dysphoria groups. Evidently, the only symptom that was clearly associated with chronic dysphoria was body-image dissatisfaction.

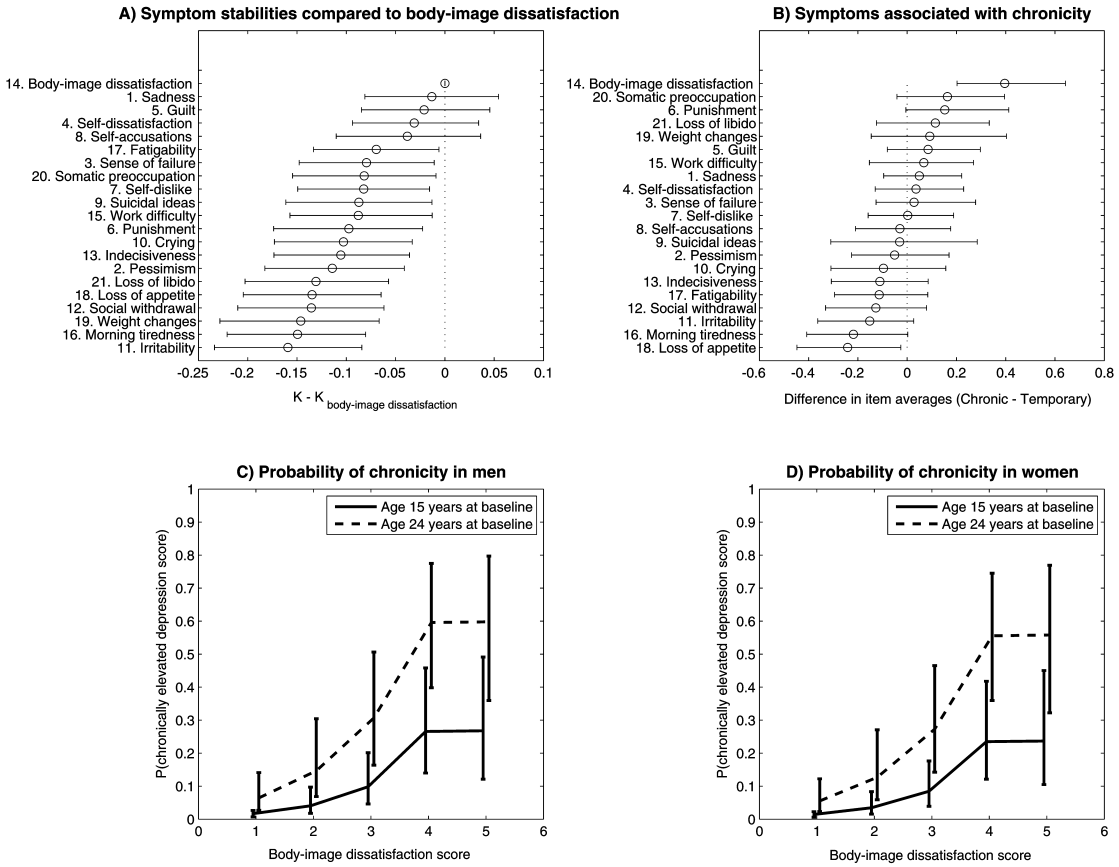

Figure 5. Symptom stabilities and symptoms associated with chronicity. (A) Symptom stabilities (parameter *K*) compared to temporal stability of body-image dissatisfaction ($K_{body\text{-}image\ dissatisfaction}$). Whiskers denote 95% bootstrap confidence intervals. B) Differences between average symptom levels in chronically versus transiently dysphoric participants, ordered by magnitude. C) Probability of chronic dysphoria for the least and most chronicity-prone male cohorts in Young Finns data as a function of the baseline (year 1992) body-image dissatisfaction, as estimated by a logistic regression model. (D) Same as panel C, but for the female participants.

Average body-image dissatisfaction was 0.40 points higher in chronic than in the transient group, where 4 = 5 - 1 points is the absolute maximum possible difference. In a permutation test, the *p*-value for no group difference was $3 \cdot 10^{-4}$, which easily withstood a conservative Bonferroni correction (Abdi, 2007) for the 21 comparisons (corrected $p = 0.006$). A very similar result was obtained in a sensitivity analysis where only 15 symptoms closest to the symptoms in DSM-IV depressive-episode definition were used in the definition of chronic and transient dysphoria groups, these definitions not including the "body-image dissatisfaction" item (not shown here, but see online supplement of the original contribution II). The average Cohen's effect size *d* for the body-image dissatisfaction score between chronic and transient groups was 0.35. Although 0.3 is generally labeled a "small" effect and 0.5 as a "medium" effect, Cohen emphasized that context needs to be considered in interpretation (Cohen, 1988): the effect 0.35 was 1.49 times the second largest absolute effect, 3.43 times the average, and 4.31 times the median absolute effect size for the group differences among the 21 symptoms, and the variance in mBDI was restricted to high scores to begin with.

Regarding predictive and causal interests, it is relevant to know what can be said about future chronicity given a present observation of a risk symptom. As body-image dissatisfaction was recognized above as a risk symptom, the chronic and transient dysphoria groups were defined otherwise similarly as before but by excluding the body-image item itself from the definitions. Then levels of body-image dissatisfaction were encoded to binary 'dummy' variables (intercept for score 1 and indicator variables for scores 2 through 5) that were used in an age- and cohort-adjusted logistic regression model. The advantage of the dummy-coding as opposed to linear encoding with a single body-image dissatisfaction score is that this way a potential dose-response relationship is revealed by the data, rather than assumed or forced by the model. Figures 5C and 5D present the results, which indeed revealed a dose-response relationship.

As evident by comparing Figures 5C and 5D, gender played a negligible role in predicting chronicity, despite women having 0.26 standard deviations higher average levels of body-image dissatisfaction than men, and higher average mBDI. The fact that high body-image dissatisfaction implied chronic dysphoria with a higher certainty for

older participants (24 years at baseline) than for younger participants (15 at baseline) may relate to generally higher levels of body-image dissatisfaction in youth (not shown here; supplementary online analysis in original contribution II). Altogether, the association between high body-image dissatisfaction at baseline and chronic dysphoria in the following 16 years was quite strong, with the probability of chronic dysphoria being 22.3 to 53.2 percentage units higher for those with most appearance pressures (score 5 in the item) compared to those with least pressures (score 1). But if chronic dysphoria is most strongly associated with the body-image dissatisfaction, what symptoms might predict more transient changes in depression?

Morning tiredness and irritability were some of the least stable symptoms (see Figures 4 and 5A), and the Introduction section discussed that much research effort has been put into elucidating causality between sleep problems and depression (sum scores). Hence, analysis of sleep-related issues holds promise for understanding more transient changes in depression. Because less prospective sleep-problem measurements existed in Young Finns data than prospective depression measurements, and because temporality between sleep and depression may be confounded by sleep problems being a prodromal symptom (Perlis et al., 1997), the situation of short-term predictive symptoms was seen as a fruitful opportunity to test and apply newly introduced pairwise measures of causality. Indeed, all depression sum scores had a non-Gaussian distribution, as required by these methods, and at least mBDI exhibited a linear relationship with sleep problems, as also required by the method (Figure 2D). Furthermore, the Young Finns study included data on the socioeconomic status of participants' parents in the participants' youth and from the participants themselves in their adulthood; these questionnaire data almost certainly exhibit a causal relationship from former to latter variable rather than the other way around, thereby providing suitable real-world benchmark for further testing of the novel methods.

Several simulation studies have shown that pairwise measures of causality do function as intended, and that they are robust against measurement errors (Hyvärinen & Smith, 2013; Shimizu et al., 2011, 2006), and even against partial confounding by an unobserved third variable (see the original contribution III). Here we concentrate on

real-data results, as shown in the Table 4. First, one sees that, in the situation with (practically) known causality, all three pairwise measures of causality correctly indicated that parents' socioeconomic status causes that of their offspring's, to a significantly greater degree than for the reverse causation. In such clear situations, the statistics therefore appear to function as intended in the epidemiological questionnaire data. Second, one sees that all three statistics suggested that sleep problems cause mBDI more than the other way around, although the tanh-based statistic was quite inconsistent across the bootstrap data sets. The inconsistency in the tanh-, or kurtosis-based, estimate may be due to mBDI being close to the Normal-distribution kurtosis; for this kurtosis-based statistic then, the non-Gaussianity requirement of LiNGAM is not as well satisfied as for the other statistics. For the skewness-based statistic, the results were more consistent with the general DirectLiNGAM-based statistic, and skewnesses of mBDI measures consistently differed from the Normal-distribution skewness (see Figure 2A and the first paragraph of section 4.1). The relationship between mBDI and sleep problems was quite linear, as required by the pairwise causality estimation (Figure 2D).

Although, BDI-II and mCES-D violated the required linearity assumption with respect to sleep problems, Table 4 nonetheless shows the causality statistics for them as well. One sees that opposite causality, from depression to sleep problems, is suggested for the BDI-II that provides more information on more severe symptoms than mBDI (Figure 2C). The results are less consistent for BDI-II over the bootstrap re-samples than for the mBDI, and still less consistent for mCES-D. Extreme caution is needed in interpreting causality statistics for BDI-II and mCES-D, as they violated the linearity assumption of LiNGAM, and suffered from larger attrition effects than mBDI.

This ends the symptom-level analyses section for this thesis, but it should be evident that much work remains to be done in this field. In fact, entire research programs on symptom-level analysis have been recently initiated [*e.g.*, Denny Borsboom's http://www.psychosystems.org/ and Peter de Jonge's "deconstructing depression" (de Jonge, 2012)]. Despite the aims towards "deconstruction" of depression, there is no

doubt that the known symptoms frequently do co-occur—to the extent and prevalence that suggests some adaptive side for the symptoms.

Table 4. Pairwise Causality Comparisons for 2000 Bootstrap Re-samples

	Chosen as cause %		Summary of values	
Method/Statistic	**Parents' SES**	**Offspring's SES**	**Statistic**	**95% confidence int.**
DirectLiNGAM[b]	100.00	0.00	0.1062	(0.0627, 0.1485)
Skew-based	100.00	0.00	0.0721	(0.0454, 0.1019)
Tanh-based	99.90	0.05	0.0077	(0.0033, 0.0124)
	mBDI	**Sleep problems**		
DirectLiNGAM[a]	00.40	99.60	-0.0433	(-0.0747, -0.0090)
DirectLiNGAM[b]	01.40	98.60	-0.0354	(-0.0677, 0.0001)
Skew-based	2.80	97.20	-0.0276	(-0.0565, 0.0009)
Tanh-based	28.50	71.50	-0.0013	(-0.0054, 0.0027)
	BDI-II	**Sleep problems**		
DirectLiNGAM[a]	77.65	22.35	0.0213	(-0.0332, 0.0781)
DirectLiNGAM[b]	100.00	0.00	0.1633	(0.0927, 0.2572)
Skew-based	100.00	0.00	0.0913	(0.0457, 0.1507)
Tanh-based	65.95	34.05	0.0011	(-0.0038, 0.0058)
	mCES-D	**Sleep problems**		
DirectLiNGAM[a]	0.00	100.00	-0.8798	(-0.8940, -0.7940)
DirectLiNGAM[b]	0.00	100.00	-0.5655	(-0.6031, -0.5185)
Skew-based	100.00	0.00	0.0443	(0.0205, 0.0730)
Tanh-based	99.85	0.15	0.0042	(0.0013, 0.0071)

Note: a) Non-standardized original variables (not available for SES); b) Standardized z-score variables; skew- and tanh-based statistic always require standardization. Second and third column report the percentages of 'wins' in the indicated pairwise comparison, whereas the two last columns summarize the statistic implying the result over the 2000 re-samples. "SES" = socioeconomic status, "mBDI" = modified Beck's Depression Inventory; "BDI-II" = Beck's Depression Inventory II, "mCES-D" = modified Center for Epidemiologic Studies Depression scale.

4.3 Mathematical results from the dynamical model

When behavioral modes of cooperation, defection, ('altruistic') punishing, and nonparticipation follow social imitation-based dynamics, the prevalances of applied strategies vary in time as a function of prevalences of the other strategies. The average time that a population spends in a cooperative mode of behavior (including the altruistic

Punishers) is a crucial factor regarding the net gains (public goods) for the group engaged in joint enterprises. Figure 6 shows the average system state (stationary distributions of strategies) in the public good game of section 3.3.1 as a function of imitation-selection strength s_i; the lower-right panel also shows the average productivity of the joint enterprise in a group with all four strategies available, in a group that lacks the Punisher strategy, and in a group that lacks the Nonpartipants. Parameters were set as in previous work, and obtaining the exact same results served to ensure that the present model-implementation concurs with the previous ones (Hauert et al., 2007; Sigmund, 2010), and the results are also useful for illustrative purposes. When the joint enterprise was irrelevant for strategy imitation ($s_i = 0$), all strategies were equally frequently observed in all models. When s_i grew, some strategies were preferred by individuals within a group more often than other strategies. When all four strategies (Cooperators, Defectors, Nonparticipants, and Punishers) were available in a group, the conditionally cooperative Punishing strategy was found most often within the group. When punishing strategy was lacking, 'rock-paper-scissor' dynamics settled in, and the three remaining strategies oscillated with more even proportions. When the nonparticipating strategy was lacking, Defectors took over the population for moderate to large s_i. For a very small s_i, when the joint enterprise was not very significant for individual participants, income of the joint enterprise was largest for three-strategy game including Cooperators, Defectors, and Punishers. When s_i grew, however, the importance of the option to abstain was increasingly reflected in the average income, and therefore in the fitness (in the parameters F and G), of the groups.

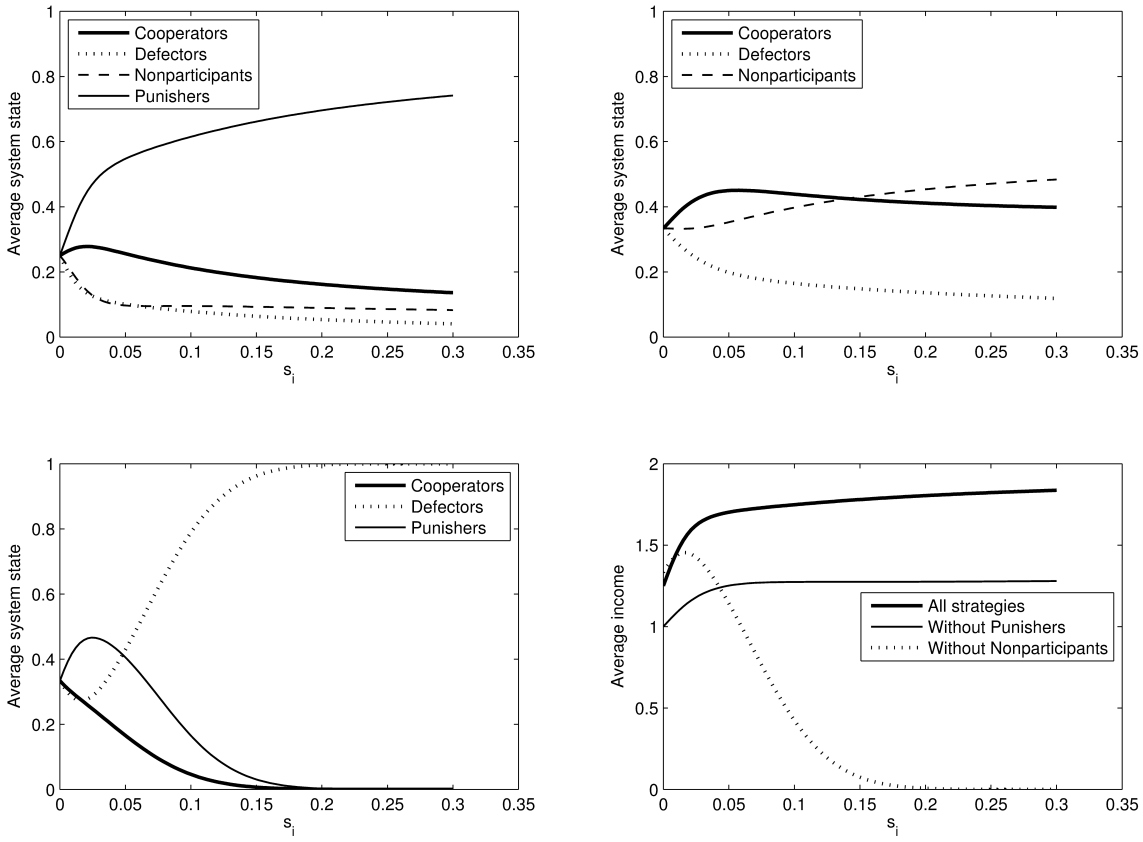

Figure 6. Average system states (stationary distribution, or relative time in homogenous states of one strategy) as a function of selection strength (s_i) and available strategies (panel legends). The lower-right panel shows the average income of the modeled joint enterprise, or public good game, for all three settings of available strategies (see other panels). Model parameters were set to equal those in the previous study by Hauert and others (2007); that is, $M = 100$, $r = 3$, $c = 1$, $\sigma = 1$, $\gamma = 0.3$, $\beta = 1$, and $N = 5$.

Analyzing the setting without Punishers, where the infinite group-size analysis is stable, provides some insight to the role of Nonparticipants; namely, nonparticipation is a response to defection (Fowler, 2005; Hauert et al., 2007; Sigmund, 2010). When 'unfair' sharing of spoils of the joint enterprise increases in a group, this is followed by an increasing prevalence of nonparticipation, which can then be followed by cooperation, that again invites defection, and the cycle starts anew—the game is locked into a fruitless rock-paper-scissors dynamics. In fact, Sigmund provided an analytic calculation showing that, in an infinite population without Punishers, the time-average

of total income indeed equals the Nonparticipants' σ (Sigmund, 2010). In order to provide intuition for the reader, and to support the following Discussion section, the dynamics implied by Sigmund's equations are visually illustrated in Figure 7.

Finally, Figure 8 shows the probability of all groups (entire population) having a Nonparticipation strategy (a depression-vulnerable genotype) in their pool of available strategies; that is, the figure shows the average time spent in a state where genetic variants for depression exist, as opposed to a state where they do not exist, provided that the incomes of a group's joint enterprises are reflected upon that group's fitness. The probability of genetic variants for depression (at a randomly chosen time) is plotted as a function of the importance of the joint-enterprise outcome for individuals (imitation-selection strength s_i in social dynamics) and importance of outcome for the group's survival (genetic selection strength s_g among groups/species/tribes/etc.). The results show that, for most selection-strength values, the probability of genetic variants for depression grows as a function of both the number of competing groups and the number of individuals per group; with sufficiently many groups and sufficiently many individuals in them, depression genetype is a certainty imposed by natural selection under the modeled dynamics. When the importance of joint enterprises is very low for the involved individuals ($s_i \leq 0.006$) but high for the groups, the pay-off can be higher for groups without the Nonparticipants (Figure 6), implying that selection favors groups without genetic variants for depression (Figure 8).

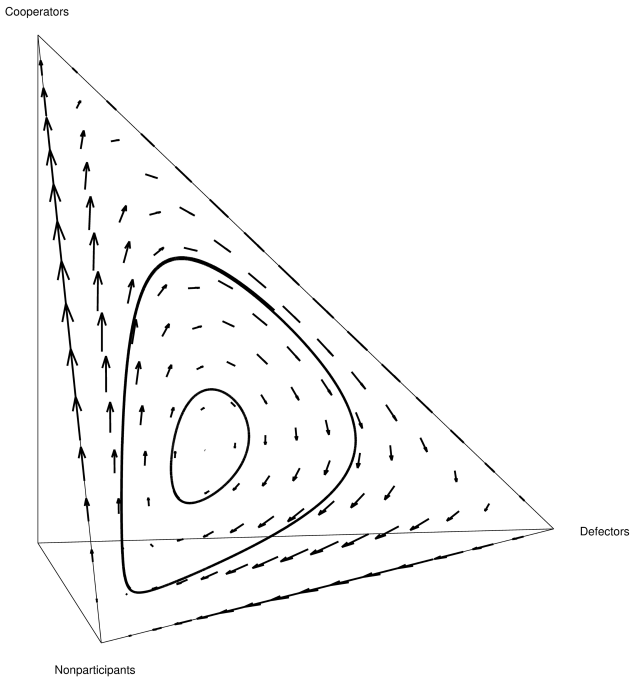

Figure 7. Time-dynamics of a joint enterprise in an infinite population with only three strategies available: defecting ('y-axis' in right-handed coordinates), cooperation (z-axis), and not participating (x-axis). All possible configurations for the proportions of strategies applied in the population are contained in a simplex (front triangle of the geometric illustration); corners describe a population with only one strategy, whereas the middle-section corresponds to the population where all the strategies are present in equal amounts. The arrows in the figure show the direction of more lucrative strategies for an individual situated in a population with the amounts of strategies indicated by the base of an arrow. The exact time-dynamics of a joint enterprise depend on the pay-off structure of the enterprise, but all enterprises that are (potentially) more beneficial than Nonparticipation and fulfill the condition $r > 2$, show similar oscillating 'Rock-Paper-Scissors' dynamics: depending on the initial condition, state of the population follows some closed orbit (thick lines provide examples) on the direction determined by the gradient field (arrows). The time-average for pay-off of the joint enterprise is always equal to that gained without participating; due to incentive to defect, the joint enterprise does not benefit the population in the long-term unless additional strategies are available. If the gain of the enterprise is small (the parameter r satisfies $1 < r \leq 2$), instead of oscillations, all possible trajectories eventually lead to Nonparticipants taking over the population. The figure was produced using equations by Sigmund (2010), with parameter-values $r = 3$, $c = 1$, $\sigma = 1$, and $N = 5$; as in the finite population model, and in other figures.

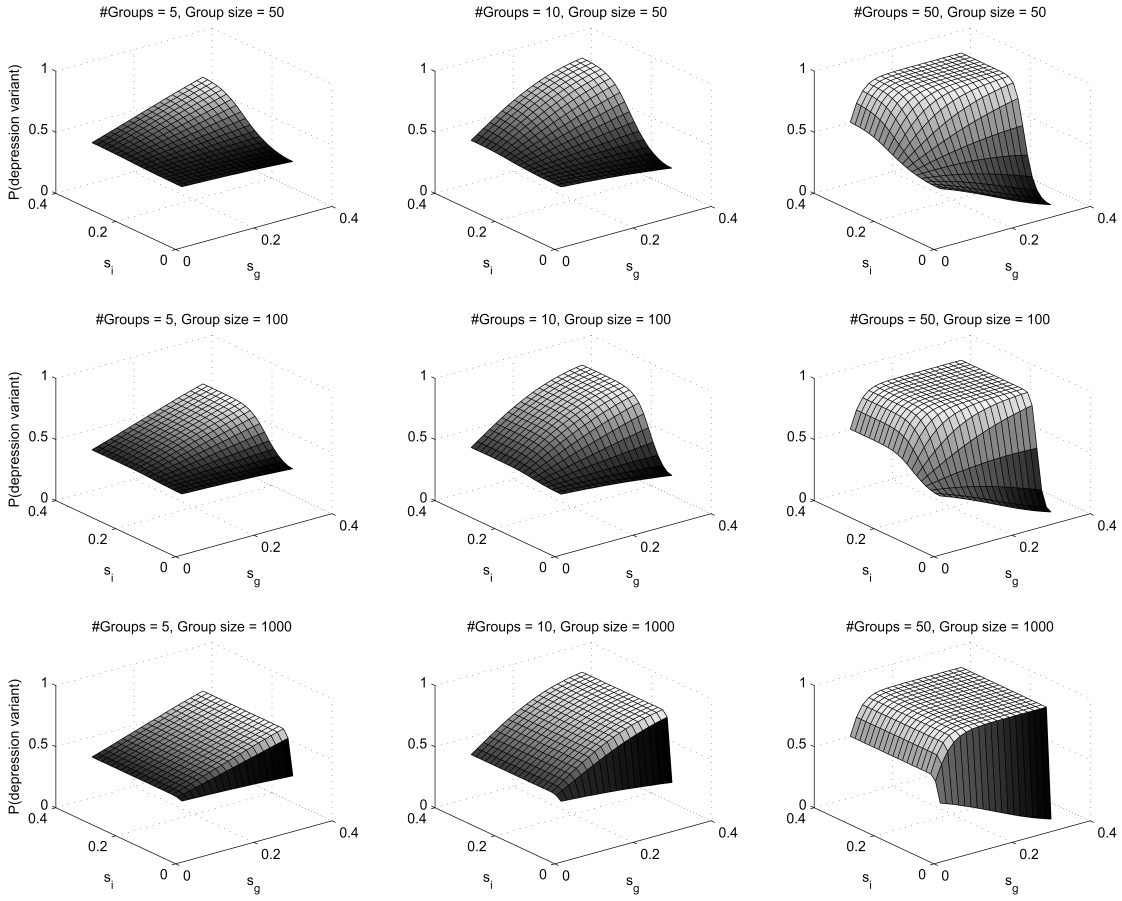

Figure 8. Probability of the system state with genetic variants for depression, or a nonparticipation strategy (stationary distribution of the state). Probability of depression genotype is shown as a function of imitation-selection strength s_i, genetic-selection strength s_g, number of groups ("#Groups"; columns of panels), and individuals per group ("Group size"; rows of panels). Other model parameters were set to equal those in Figures 6 and 7, and in the previous study by Hauert and others (2007); that is, $r = 3$, $c = 1$, $\sigma = 1$, $\gamma = 0.3$, $\beta = 1$, and $N = 5$.

5 DISCUSSION

This thesis studied depressive symptoms in general populations. The study was divided to empirical and theoretical approaches, and the former were further divided to the study of widely used sum scores and to the study of individual symptoms. An empirically justified statistical model for the trajectories of sums of depressive symptoms in the general population was estimated. Regarding individual symptoms, body-image dissatisfaction was recognized as being more stable in time than the other depressive symptoms, and more strongly associated with chronically elevated levels of symptoms than any other examined single symptom. As a fully theoretical contribution, bargaining models of depression were united with a dynamical model for evolution of cooperation, resulting in a more comprehensive adaptive theory of depression compared to the frequently criticized previous attempts. In the following, these findings are first discussed in more detail (sections 5.1-5.3), and then brought together (sections 5.4-5.6).

5.1 On the time trajectories of depressive-symptom sums

As one of its aims, this thesis examined the longitudinal course of a depressive-symptom sum in adulthood using the Young Finns follow-up data (Raitakari et al., 2008). Initially fifteen to thirty year-old participants were followed for sixteen years. Four successive measurements were taken, with five, four, and seven year intervals between. A continuous-time stochastic differential equation model was estimated for the temporal progression of sum scores; possibly for the first time in epidemiological depression research. The observed population variance in the temporal course of depressive symptoms was decomposed into the three main sources: (1) individual differences in stable trait-like depression that may be caused by genetic differences, or by early experiences that determined the subsequently time-constant levels of depressive symptoms for the participants; (2) individual differences that derive from a continuously accumulating trajectory of small changes, such as negative (or protective)

cycles of deteriorating (or improving) mental health that influence the course of depression, and from the differential temporal behavior of different symptoms; and (3) differences due to state fluctuation or measurement error without continuity over time. In men, the stable trait accounted for the majority (61%) of the total variance, followed by cumulative effects (23%), and state fluctuations (16%; Figure 3A). In women, the cumulative sources were more important than among men and accounted for 44% of the variance, followed by stable individual differences (32%) and state fluctuations (24%). In addition to the differences in the longitudinal course, women had a higher mean-level and higher variance of depressive symptoms than men, in all follow-ups.

The present data and model estimates were in agreement with the previous observations that women have higher mean-levels and higher variance of depressive symptoms than men; this gender difference emerges by ages 13 to 15, which is before the first follow-up in the present data (Nolen-Hoeksema & Girgus, 1994; Hankin et al., 1998; Hyde et al., 2008). During the age years followed here, however, a declining trend in population-average of depressive symptoms has been typically observed (Galambos et al., 2006; Merikangas et al., 2003). The present model estimates also found this declining trend for both men and women alike (Figure 3C); women had higher levels of depression throughout the 16 year follow-up, but also a steeper decline, than men.

In both sexes, however, temporal changes in affect generally are not unconstrained, but eventually tend to return to an individual-specific "affective home base" (Clark et al., 2008; Kendler et al., 1998; Kuppens, Oravecz, & Tuerlinckx, 2010). In the "Dynamics of Affect" (DynAffect) model of Kuppens, Oravecz, and Tuerlinckx (2010), fluctuations in an individual's affective state are described by Ornstein-Uhlenbeck process with individual-specific parameters, plus state/measurement-error variations. Since the EDM model consists of an Ornstein-Uhlenbeck process with individual specific trait-levels, or "affective home base", superimposed to state variations, the EDM is very close to the DynAffect model. Although the DynAffect is a two-dimensional model for affective valence and arousal (Kuppens et al., 2010), and the present EDM was one-dimensional model for dysphoria, probably the most important difference is the following one. Here the linear coefficient A, which is negative of the

"attractor strength" β in DynAffect (Kuppens et al., 2010) and in Ornstein-Uhlenbeck process (Oravecz et al., 2011), was modeled separately for men and women, whereas the DynAffect models it separately for each participant. Therefore, the DynAffect is a much more flexible model than EDM. Most likely, too flexible for epidemiological panel data with few prospective measurements; although, the issue might be worth a future analytic examination. The present study had complementary sides compared to the DynAffect study.

Whereas Kuppens and others (2010) followed 80 or 60 participants accurately over periods of two weeks or 4 days, the present study followed less accurately 3061 participants over a period of 16 years. Kuppens and others inferred that "core affect may to some extent be self-similar across time scales", because the same (*i.e.*, Ornstein-Uhlenbeck) process governed affect trajectories of two different sampling frequencies [hours/days/weeks *vs.* minutes (Kuppens et al., 2010)]; a conclusion further supported by the present observations about similar processes over the course of 16 year dysphoria trajectories. Whereas Kuppens and others (2010) studied variability in attractor strengths (*c.f.*, random-effect variance for A), they did not specifically study gender differences. The present study suggests that gender differences matter, and would be interesting to study also within the DynAffect framework. Comparing the present study and DynAffect is, of course, complicated by the question of the degree of correspondence between the simple measure of pleasure and arousal applied by Kuppens and others (2010) *versus* the sum of 21 depressive symptoms studied herein.

Previous studies have rarely decomposed the non-stable variance in depression into a state/error component versus a component that continuously develops over time. Here, it was estimated that less than 27% percent of variance is due to state-like fluctuations or measurement error. The rest is accounted for by autoregressive trajectories of depressive symptoms that accumulate over time differently for different participants, or by trait-like depression that stays constant for an individual, with between-individual variation in that constant. The latter share of the between-individual variance is either constitutional or has stabilized before adulthood. The cumulative part must be due to changes in individual or in environment that have occurred during the follow-up period

of sixteen years. Genetic effects can potentially enter the cumulative part only via gene-environment interactions (Caspi, Hariri, Holmes, Uher, & Moffitt, 2010), via pre-programmed lagged effects, or via epigenetic effects (T. Zhang & Meaney, 2010), but not as solid constitutional genetic effects. The cumulative proportion of intra-individual variation is not measurement error, yet reflects changes in time.

In the current research literature on depression, the cumulative source of variation is the least well-known among the three components describing the longitudinal course of depressive symptoms. Recall that we modeled an individual's time accumulating depression-level using integration with respect to Brownian motion (a.k.a. the Wiener process). Brownian motion is the random process that is most frequently utilized in the modeling of time trajectories of agents or objects that are continuously perturbed by many small outside influences (Klenke, 2008; Øksendal, 2003). For example, it is a model for a very small particle immersed in fluid and undergoing movement due to countless tiny collisions with the surrounding moving atoms. Instead of moving atoms, we modeled effects due to many small external and internal occurrences affecting depression and adding up in time (also the perturbations due to the naïve summation of differentially behaving symptoms are included in the modeled random "collisions"). The fact that women were estimated to have more such occurrences is in line with the multiple pathways model that aims to explain the gender differences in depression (Hyde et al., 2008). The modeled cumulative source of variance can also be motivated by the finding that differences in depression and anxiety levels between twins grow as they age, until saturating in old age (Kendler et al., 2011). A Brownian motion-based model is also compatible with the concepts of multi- and equi-finality (Cicchetti & Rogosh, 1996), because the same starting value of depression can develop to different outcomes at a later time; and different initial values can converge. In effect, unobserved influences may or may not cancel each other out as time passes; although, there is an average tendency for individuals to drift towards their base-levels of affect.

In addition to the DynAffect and multiple pathways models, the model of time-accumulation that was applied here is closely related to an autoregressive time-series approach that was utilized in a prior study of children and adolescent depression, along

with the trait and state components (Cole & Martin, 2005; Kenny & Zautra, 1995). As seen from Figure 1, however, the present continuous-time model can be more readily compared over studies with different depression assessment-time intervals than the discrete-time autoregressive models (Oud & Delsing, 2010; Oud & Jansen, 2000). Using the parameter estimates in the original contribution I, various standard predictive calculations become possible for depression trajectories, including the expected number of level-crossings (*c.f.*, episodes), and the population mean and variance trajectories (Klenke, 2008; Øksendal, 2003). As a word of caution, however, limitations existed in the parameter estimation accuracy (section 5.5).

5.2 On the differential behavior of individual symptoms

In addition to trajectories of sums of depressive symptoms, this thesis studied the differential behavior of the individual symptoms. More specifically, stability of symptoms, symptoms associated with chronic dysphoria, and causality between sleep problems and depressive-symptom sums were studied. Perhaps the most striking finding of these explorations was that body-image dissatisfaction was the most temporally stable depressive symptom in the mBDI inventory, and clearly associated with chronically elevated sum score, or chronic dysphoria. Depending on age and birth cohort, having expressed most intense (score 5) fear of "looking ugly and displeasing" at the baseline implied from 22 to 53 percentage units higher probability for chronic dysphoria over the 16 year follow-up compared to those with no appearance pressures (score 1). Regarding more transient changes, novel causality algorithms supported the notion that dysphoria in general population is more likely to be caused by sleep problems than to cause them. Irritability and morning tiredness were among the least temporally stable symptoms, suggesting that sleep problems may cause more *temporary* dysphoria than body-image dissatisfaction. What is theoretically noteworthy in these results is that they concur with the viewpoint where symptoms are a part of a more general causal network (Cramer et al., 2012, 2010) than the frequently applied one

where the depressive symptoms are (implicitly or explicitly) interpreted as reflections from a latent cause (Borsboom et al., 2003; Cramer et al., 2012; Reise & Waller, 2009).

Although average body-image dissatisfaction was higher in youth than in later adulthood, the between-individual differences in it were very stable, signifying the importance of the relative position among ones age group. Less strong association between body-image dissatisfaction and chronicity in younger cohorts (Figures 5C and 5D) may nonetheless relate to puberty being tumultuous for many, whereas adulthood perceptions may have a momentum of personal history behind; possibly representing a true state leading to ostracizing by others (a depressing event), or alternatively, a developed dysfunctional attitude towards oneself (Beck, 1967, 2008). In the cognitive model of depression, dysfunctional attitudes towards oneself induce a cognitive vulnerability that can be triggered by stressful life-events; a process that may culminate in pervasive cognitive biases and depression (Beck, 1967, 2008; Disner, Beevers, Haigh, & Beck, 2011). A translation of the "body-image dissatisfaction" item reads "I fear that I look ugly and displeasing". The fact that a participant strongly endorses such a statement about him- or herself in a formal questionnaire could plausibly reflect a self-defeating cognitive bias. Furthermore, a recent meta-analysis indicated that current media coverage serves to induce dissatisfaction towards ones physical appearances (Grabe, Ward, & Hyde, 2008). Hence, the present finding can be interpreted in terms of the cognitive model of depression. Also, when comparing cognitive-behavioral therapies for major psychiatric disorders, they appear to work best for the unipolar depressive disorder (Lynch, Laws, & McKenna, 2010).

Body-image dissatisfaction may also predate the formation of cognitive vulnerability, however. Such a view point is consistent with another, evolutionary, theory of dysphoria. It has been suggested that depressed and low moods may function as a learning and motivational signal that is analogous to a physical pain signal but for the context of social interaction (Hagen, 2011; MacDonald & Leary, 2005; Nesse, 1991). For example, MacDonald and Leary (2005) suggested that adaptations for responding to social exclusion evolved on top of the existing biological capacities for experiencing pain signals, and that similar to physical pain, social pain can also become chronic. In

that theory, the most important trigger of social pain is the threat of social exclusion. Since physical appearances are associated with signals of social acceptance versus rejection, the fear of looking ugly and displeasing may elicit social pain as a response to the comparatively high probability of social exclusion. Alternatively, depression could arise as a bargaining reaction (Hagen, 2003) to appearance-related discrimination (Harper, 2000). Because looks are more difficult to change than behavior in response to the exclusion threat, a chronic social pain and associated dysphoria may ensue. Therefore, also prediction of several evolutionary hypotheses would be that "body-image dissatisfaction" associates with the chronic dysphoria, as was found in the present study. These hypotheses seem to also fit better than the cognitive model with the finding that depressed persons blamed others about their negative life events 8.5 times more often than they blamed themselves (P. Gilbert, J. Gilbert, & Irons, 2004).

In contrast to the temporally stable symptom of body-image dissatisfaction, symptoms like irritability and morning tiredness were the least stable ones. These mBDI symptoms were closest in content to sleep problems, and several studies have tried to resolve the causal direction between sleep problems and depression. Some studies have found sleep problems to precede dysphoria or depression (Almeida et al., 2011; Paunio et al., 2009), whereas other studies also report the opposite temporality (Jansson-Fröjmark & Lindblom, 2008; Sivertsen et al., 2012), and the causality matter is further complicated by suggestions that sleep problems might be a prodromal symptom (Perlis et al., 1997). That is, both sleep problems and other depressive symptoms could be responses to an unobserved single causal antecedent, but one or other might be a faster response (prodrome refers to an early symptom). Network theorists suggest, however, that instead of postulating latent variables, one should first consider frequencies of commonsensical direct causal paths of symptoms, such as "insomnia → fatigue → concentration problems → self-reproach [due to bad performance] → depressed mood" (Cramer et al., 2012). In line with this latter viewpoint, this thesis studied whether the two-dimensional population distributions of sleep problems and sums of other depressive symptoms reveal indications of dominant causality to one or other direction. Hence, recently introduced causality estimators, that are able to estimate causality from

cross-sectional data (Hyvärinen & Smith, 2013; Shimizu et al., 2011, 2006), were tested on this epidemiological problem that can be considered genuinely open with respect to the issue of causality.

It was first shown that each of the three applied causal estimators easily recognized the correct causality from a benchmark data consisting of parents' and their offspring's socioeconomic status. Then sleep problems and depression was analyzed in two different data sets with three different depression inventories. The estimators quite consistently indicated that sleep problems caused depressive symptoms when mBDI, that best fulfilled the required assumptions for causality estimation, was studied. In the same Young Finns data, another depression measure (BDI-II) yielded an inconsistent result, but this data set violated the assumptions of the model. In addition, the mCES-D inventory in the Wisconsin Longitudinal Study's data also violated the assumptions, and also provided conflicting results among the different estimators. The violations of linearity assumption (BDI-II and mCES-D) appeared to result from a floor-effect in the two inventories compared to sleep problems; for example, BDI-II was informative on more severe depression than mBDI, but most participants' statuses fell short from its optimal range, being best represented by mBDI (see Figure 2B and 2C, and the related results). Also, sleep complaints are quite common nowadays, and were linearly related with mBDI, showing an opposite quadratic effect with BDI-II compared to that between mBDI and BDI-II.

Hence, in the single reasonably reliable case, sleep problems appeared as more frequent causal antecedent of the sum of other depressive symptoms than *vice versa*. As indicated by the expectation operations in the methods description, the results from causality algorithms refer to dominant causality in the population, not to a mechanistic or certain relationship (also simulations in original contribution III demonstrate that the methods are robust for situations where a minor part of the population has a different relationship). Nonetheless, the present result was for observations of less severe depression than in many clinical samples, and it is quite possible that severe depression has a greater role in deteriorating sleep quality relative to less severe depression. The present results therefore suggest that causal paths starting from sleep problems and

ending in generalized dysphoria are more frequent than the opposite kinds of causal paths among the general population. Cross-sectional estimation of complex paths such as "insomnia → fatigue → concentration problems → self-reproach → depressed mood" was not attempted, as reliable estimation at such a level of detail is likely to require accurate prospective sampling; cross-sectional causality estimation is an assumption-heavy, and therefore error-prone, business. Yet, temporality can be confounded, experimental studies can be unethical, and dense temporal sampling is often expensive and may involve unnatural repetition of questionnaires. Hence, the cross-sectional causality estimation seems a promising addition to available methods for depression research, and also for general epidemiology; especially, when adopting the causal network viewpoint.

Despite interpreting depressive symptoms as a network phenomena, it may still be possible to identify common causes for wide-spread network activations. For example, network models frequently draw from statistical mechanics (Boccaletti, Latora, Moreno, Chavez, & Hwang, 2006), where many complex multi-particle interactions and phase transitions have been satisfactorily understood in terms of certain control parameters (Baxter, 2007). In biology, an astonishing complexity has been understood from the simple principle of survival of the fittest, dynamics of evolution being increasingly captured into the "equations of life" (Nowak, 2006). This thesis suggested that a key 'control parameter' of depressive symptoms is their role in instantiating a nonparticipation strategy in otherwise compulsory joint enterprises (or Nonparticipants' pay-off).

5.3 On the evolutionary origins of depressive symptoms

This study showed that when outcomes of joint enterprises drive social competition among individuals within groups *and* an evolutionary competition among a finite number of groups, evolution is likely to introduce some means for individuals to avoid forced participation to within-group joint enterprises. More precisely, if the joint

enterprise is sufficiently important for both the between-individuals and the between-groups competition, and many large groups exists, then a Nonparticipation strategy will be present almost surely (with probability one). When the enterprises are not very important for the participating individuals, free-riding can be partially overcome even without a Nonparticipation strategy. The present work interpreted depression as a strategy that has provided a way to not participate in otherwise nearly compulsory joint enterprises during evolutionary history. This interpretation was titled as the *nonparticipation hypothesis* of depression.

The analysis was not intended to belittle the individual-level suffering among the depressed, but to provide insight to evolutionary origins and group-dynamic aspects of depression. Several theories of depression have interpreted it as means for avoiding participation in joint enterprises with undesirable terms [*e.g.*, (Andrews & Thomson Jr, 2009; Hagen, 1999, 2003; Price et al., 2004; Watson & Andrews, 2002)]. Genetic variants that allow one to become depressed may have helped to ensure a nonparticipation strategy in join enterprises among groups of people during the evolutionary history, resulting in higher fitness compared to groups without such genetic variants, and therefore in further spreading of the genes. The theoretical results obtained in the present study supported this view, although they naturally depend on the modeling assumptions that were made.

The following text reviews some additional empirical evidence concurrent with the modeling of depression as a nonparticipation strategy in a public good game (*i.e.*, in joint enterprises). To begin with, an analysis without the Punishers clearly showed that Non-participation is a response to defection. Therefore, when 'unfair' sharing of spoils of a joint enterprise increases in a population, this is followed by an increasing prevalence of nonparticipation (Figure 7). If the situation is does not admit of an alternative enterprise with comparable gains, or players are forced to participate, then depression is hypothesized to emerge as a substitute for nonparticipation. In general, a positive correlation between unfair social situations and depression is expected. Such empirical correlations can be found.

Income inequalities offer one example where 'unfair' sharing may be involved, and clear Punishing strategy may not be available: everyone is required to participate in the joint enterprise of economy (other 'nonparticipation' strategies are not readily available), high income inequality implies that some ('Defectors') benefit from the work of others ('Cooperators'), and making profit is often legally accepted, even encouraged (no 'Punishers'). Therefore, imitation dynamics predicts that disproportionate increases in income inequality are followed by increases in depression-based nonparticipation. Indeed, income inequality is more clearly associated with mental health than with the physical aspects of health (Muramatsu, 2003).

Village-level income inequality among foraging-farming societies in the Bolivian Amazon (Tsimane') is associated with more negative emotions, independently of social capital, residential segregation, and public policies (Godoy et al., 2006). Many other studies also report associations between income inequality and depression or mental health (Fiscella & Franks, 2000; Weich et al., 2001). Depression increases work absenteeism and early retirement, and decreases productivity (Benden & Farvolden, 2010; Karpansalo et al., 2005). Therefore, 'unfair' conditions (presence of defection) seem to promote depression, and the depressed do not fully participate in the joint enterprises of the society. Generally, unequal returns for equal efforts lead to avoidance of the joint enterprise in question; even monkeys refuse to participate if they witness that other monkeys obtain a more attractive reward for equal effort (Brosnan & De Waal, 2003).

A study from Britain found that income inequality was more strongly associated with worse mental health in those with higher absolute income levels (Weich et al., 2001), which may explain null findings in a study that oversampled poor individuals (Sturm & Gresenz, 2002). As countries get richer, rates of mental illness increase; prevalence of mental illness being correlated with income inequality (Pickett, James, & Wilkinson, 2006). In the current modeling context this concurs with the following observation: the lower the Nonparticipants' pay-off, the more tempting the other strategies (Sigmund, 2010). Or in terms of Bargaining Model, when entering to bargaining is not just costly but *too* costly, it is avoided.

In addition to income data, work-life in general abounds with joint enterprises. Work conditions where people experience that they gain little rewards (money, esteem, promotion prospects, and job security) compared to efforts they have invested lead to lower mental health later on [depression, anxiety, chronic fatigue, and psychotropic drug consumption (Godin, Kittel, Coppieters, & Siegrist, 2005)]. A quasi-experiment has shown that increased job demands is a causal antecedent for work-abstinence due to a depression diagnosis (Kivimäki et al., 2010). In general, however, the social 'fair-play' characteristics of a work-unit appear to be more relevant for depression than work-content characteristics like job demands and decision latitude (Netterstrøm et al., 2008; Ylipaavalniemi et al., 2005).

Evidence for genetic vulnerability to depression exist (Bienvenu et al., 2011; Cross-disorder Group of the Psychiatric Genomics Consortium, 2013a, 2013b). According to the nonparticipation hypothesis, however, a significant part of the variance in depressive symptoms should derive from social imitation dynamics. Twin studies find depression to be 31%-42% heritable (Bienvenu et al., 2011; Sullivan et al., 2000), but also show that twins became less alike with respect to depression as they age (Kendler et al., 2011). This study found support for time-accumulating sources of variance (Figure 3A), and it has been estimated that more than half of the patients without any treatment nonetheless remit from clinical depression within a year (Whiteford et al., 2012). Also some controversial evidence for a genetic sensitivity to depressogenic environments exists (Caspi et al., 2010; Clarke, Flint, Attwood, & Munafò, 2010; Karg, Burmeister, Shedden, & Sen, 2011; Keltikangas-Järvinen & Jokela, 2010), as expected in gene-culture multilevel selection. Hence, we do have evidence on temporal change. But is there more direct evidence that people imitate behavioral strategies of other people, and does this imitation generalize to long-term emotional states?

In experimental public good games, cooperative and defecting behavior is imitated by others over several rounds of the game with different and unrelated participants than the origin of the behavior (Fowler & Christakis, 2010). Therefore, imitation dynamics exist in experimental public good games with human participants. Also experimental results in wild vervet monkeys show that social imitation is a potent force shaping

group differences in foraging decisions (van de Waal, Borgeaud, & Whiten, 2013), suggesting a long-term phylogenetic role for behavioral imitation in primates. The observational data from the Framingham Heart Study indicates that both happy and depressed mood spreads from one human participant to another through their social connections (Fowler & Christakis, 2008; Rosenquist et al., 2010). Confounders and homophily (tendency to seek like-minded company) were controlled for, and the results suggested that depressed mood was causally inducted from one person to another (Rosenquist et al., 2010). Indeed, happy and depressed emotions in the Framingham data can be modeled using models for infectious diseases (A. L. Hill et al., 2010). Hence, there is plenty of evidence that both people and their evolutionary ancestors do imitate each other, and that such activity is emotion-laden.

The empirically oriented Framingham studies of social dynamics leave open the mechanism of emotional 'infection', or induction (A. L. Hill et al., 2010). They also give rise to a puzzle: why (after controlling for homophily and confounders) significantly less mood induction is observed between spouses than between friends over the course of time (Fowler & Christakis, 2008; Rosenquist et al., 2010)? In the context of the nonparticipation hypothesis, these findings can be explained: one of the mechanisms of induction would be that the participants adopt their close ones interpretations' regarding the reward value of relevant joint enterprises. This can happen, for example, via direct observation of mood or via conversations. The amount of time spent with a spouse, or closer emotional connection to the spouse in comparison to friends, need not be directly relevant for many of the important joint enterprises that one is engaged with.

Experimental evidence from the context of public good games shows that a possibility to pay in order to punish co-players increases the pay-off in the long run (Fehr & Gächter, 2000; Gächter, Renner, & Sefton, 2008), and that altruistic punishment occurs in various different human populations in different continents, nations, environments, and economies (Henrich et al., 2006). These experiments do not explicitly assess the option to abstain, but surely none are forced to participate in them, as it would be against the ethical standards in human experiments. Although it is

difficult to directly assess the role of depression-related nonparticipation in naturally occurring joint enterprises, some relevant evidence exists. More depressed individuals had fewer intentions to cooperate than less depressed individuals in a game that replicated the classical Prisoner's Dilemma pay-off structure in socially salient contexts (Surbey & McNally, 1997; Surbey, 2011). Depression both results in absenteeism from a joint enterprise (Benden & Farvolden, 2010; Karpansalo et al., 2005) and is negatively associated with interpersonal trust (Lester & Gatto, 1990). Depression is also associated with self-reported lack of self-directedness, high harm avoidance, and low cooperativeness (Farmer et al., 2003; Hansenne et al., 1999; Jylhä et al., 2011; Marijnissen, Tuinier, Sijben, & Verhoeven, 2002). These behavioral tendencies serve to promote abstaining from collaborative efforts, and depressed individuals also have less links in social networks than the non-depressed (Rosenquist et al., 2010).

So far we have discussed about evidence for the role of depression in social dynamics, and that this role should imply the existence of a biological evolutionary basis; does direct evidence exist regarding a connection between physiological processes and social processes in the context of joint enterprises? The genetic basis of neurotransmitter serotonin may interact with external environment in the etiology of depression (Caspi et al., 2010; Clarke et al., 2010; Karg et al., 2011; Keltikangas-Järvinen & Jokela, 2010), but direct pharmacological manipulation of serotonin levels does not seem to provide very efficient nor immediate general intervention to depression, save maybe the acute and severe cases (Andrews, Thomson, Amstadter, & Neale, 2012; Hagen, 2011; Kirsch et al., 2008). Instead, recent experimental evidence suggests that serotonin modulates social decision-making and the associated brain areas (Crockett, 2009), and also twin-studies have found heritable variance in experimental socioeconomic-game choices (Cesarini et al., 2009).

In addition to depressed mood, the genetics of serotonin has been reported to associate with anger and hostility (Carver & Miller, 2006; Keltikangas-Järvinen & Jokela, 2010; Merjonen et al., 2011). Regarding evolution, depression has been associated with a de-escalating strategy in competition or conflict between individuals, and anger with an escalating/assertive strategy (Price et al., 2004). In the framework of

public good games, depression naturally aligns with a nonparticipating strategy, whereas anger promotes a punishing strategy. Hence, serotonin may (indirectly) modulate sensitivity to perceive social unfairness and act upon it, leaving open the choice between a nonparticipating strategy and a punishing one. The choice may depend on many other environmental, biological, and cognitive contingencies. In general, serotonin serves very many homeostatic functions, some quite peripheral to brains [e.g., gut movement or blood clotting (Andrews et al., 2012)]. While the issue is complex without doubt, viewing genetic studies through social decision-making processes and nonparticipation hypothesis could resolve many apparently conflicting findings.

The nonparticipation hypothesis explains the high co-morbidity between negative emotions (*e.g.*, depression, anxiety, hostility, and anger): they often co-occur and depend on proximal brain structures, because they are adaptations for solving the same problem—promotion of cooperation in individuals with high self-interest. Had anger evolved without the potential for depression, it might have led to populations with Cooperators, Defectors, and Punishers; shown to result in a dominance of defection (De Silva et al., 2010; Hauert et al., 2007; Sigmund, 2010). Had depression evolved without anger, populations with only Nonparticipants, Cooperators, and Defectors might have netted nothing from joint enterprises in the long-run (Sigmund, 2010). The co-existence of negative emotions serves to instantiate the four strategies that yield cooperation via voluntary participation and altruistic punishment. Indeed, participants in experimental public good games feel anger towards co-participants that fail to cooperate, and these emotions are stronger when between-group competition is present and the co-participants belong to their own group (Puurtinen & Mappes, 2009).

The nonparticipation hypothesis can also explain the therapeutic alliance effect in context of depression: in psychotherapy, alliance broadly refers to a collaborative and affective bond between therapist and patient, and its quality has been found to affect the outcome of the therapeutic process moderately, but consistently (Martin et al., 2000). Because depressed mood is physiologically and socially costly for an individual, the nonparticipation hypothesis assumed that it only arises in joint enterprises that one subjectively views as somehow unavoidable. When an alternative equally important

enterprise arises, and one takes it under, the depression should lift; this prediction could play a key role in understanding the "alliance effect".

Although exact conceptualizations of alliance vary, three central ingredients are widely agreed upon: (*i*) the patient's and therapist's ability to agree on treatment goals, (*ii*) the affective bond between the patient and therapist, and (*iii*) the collaborative nature of the relationship. Other than these three ingredients, meta-analyses have not found support for any other suggested moderators or mediators of the alliance effect (Flückiger et al., 2012; Martin et al., 2000). There may be none. When the therapist and patient enter into (*i*) an agreement on the treatment goals, they enter into an alternative joint enterprise with respect to the patient's previous and depressogenic one. The fact that they make an (*ii*) affective bond ensures that also the phylogenetically old parts of the brain register the altered socioemotional context. Finally, the (*iii*) collaborative nature of the relationship signals that this novel joint enterprise is not poisoned by defective practices, but investing in it will benefit both the patient and the therapist. Together the conditions *i* through *iii* both are the therapeutic alliance and result in lifting of the depressive symptoms through an alternative joint enterprise not requiring the evolutionary adaptation for nonparticipation (*i.e.*, depression).

The precise evaluation of the nonparticipation hypothesis is necessarily a difficult task: it is difficult to draw clear borders between numerous joint enterprises that modern individuals are immersed in, or analyze their relative importance. It is also difficult to determine the amount and quality of alternative nonparticipation strategies. Yet, some dynamic model seems necessary for a satisfactory functional understanding of depression; the present suggestion is a simple one to start with as it incorporates only four-dimensional dynamics (four strategies). Higher-dimensional dynamical systems are increasingly difficult to analyze.

A number of research implications follow from adopting a combined evolutionary- and social-dynamics view point to depression. Regarding evolutionary dynamics it is difficult to separate adaptations that promote survival of a group from those that promote reproductive fitness of an individual. Herein, it was hypothesized that it is adaptive to retain in a population those traits that prevent adverse outcomes of imitation

dynamics; namely, the evaporation of cooperation in important joint enterprises. Yet, evolutionary dynamics can also follow individual-based fitness. This is also a matter of the level of analysis, as most 'individuals' can also be seen as 'groups' of some entities; for example, humans tend to belong to several groups and hierarchies of groups, but also *contain* groups of organs that are groups of cells, and so forth. Careful analysis is needed to separate group-level effects from individual-level effects on one hand, and evolutionary dynamics from social-imitation dynamics on the other hand. Both evolutionary and imitation-based models tend to take similar mathematical forms (Nowak, 2006; Sigmund, 2010), but obviously operate on different time-scales (multiple generations *versus* life-spans of individuals). This should justify the separation of time-scales for the model of this study.

The modeling results reviewed here pertained to a 'well-mixed' population where any individual has equal changes to meet any other individual. It has been shown, however, that network structures (distribution of social contacts between individuals) can play an important role in evolutionary and imitation-dynamical models (Nowak, 2006). Therefore, future studies evaluating the non-participation hypothesis may benefit from careful assessment of the structure of social network relevant to a given joint enterprise; typically, some network structures function as an amplifier of imitation/selection, whereas some network structures attenuate imitation/selection (Nowak, 2006). Regarding present results, however, it is reassuring that a recent empirical study concluded that "population structure has little relevance as a cooperation promoter or inhibitor among humans" (Gracia-Lázaro et al., 2012).

This study considered models where punishing of others is 'costly' for the individuals participating in a joint enterprise. Punishing need not be very costly always; for example, ostracism of defectors might be easy sometimes. Yet, people do engage in punishment practices that are costly, and such behaviors seem to be an important part of evolution. The partial nature of the driving social and evolutionary forces was explicitly modeled using the selection-strength parameters (s_i and s_g). Although non-costly punishment is an interesting research question in terms of modeling the evolution of cooperation, it does not directly undermine any arguments presented here.

The nonparticipation hypothesis, in general, need not be an all-inclusive model of depression. Rather, it is intended to explain some functional aspects of depression etiology and evolutionary history. The hypothesis does not necessarily conflict with other theories of depression. For example, psychic pain associated with low mood and sadness is sometimes seen to serve functions analogous to physical pain, but in response to social rather than physical adversity (Hagen, 2011; MacDonald & Leary, 2005; Nesse, 1991). Evolved abilities often provide more than one benefit for their carriers; for an extreme example, consider the ability to grasp objects. The psychic pain hypothesis of depression does not appear exhaustive, however, as it faces difficulties in explaining suicidality and severe depression (Friedman, 2012; Hagen, 2003, 2011). The psychic pain of the depressed may also be interpreted within the nonparticipation and bargaining hypotheses. Psychic pain and pleasure may supplement or substitute the driving role of outcomes ('money') in some biologically relevant joint enterprises, providing information that one is suffering fitness costs (Table 1). Although 'money' is an intuitive conceptualization for a pay-off from joint enterprise, the models for evolutionary dynamics operate on deviations from average pay-off, or fitness, instead of any fixed units (Nowak, 2006; Sigmund, 2010).

Depressive episodes are frequently associated with stressful life events (Kendler & Gardner, 2010; Kendler et al., 1998). The causal effect of many stressful life events may be modest, however (Kendler & Gardner, 2010). Furthermore, the increase in depressive-episode risk due to temporally well-isolated stressful life events tends to be short-lived, generally disappearing in less than three months (Kendler et al., 1998; Surtees & Wainwright, 1999). Instead, the stressful life events "associated with prolonged duration of risk—as exemplified by serious financial, legal, work, and marital problems—were usually more 'difficulty-like' in that they may have more 'fuzzy' onsets and are likely to persist some time" (Kendler et al., 1998). Such "difficulty-like" events can easily be subjectively interpreted as unfair social situations, and presence of defection, which is a causal antecedent of depression according to the present nonparticipation hypothesis. Due to the cognitive paradigm, existing depression research has concentrated much on negative self-image, but an explicit study on blame

attribution for life events and difficulties by depressed persons suggested as large as 8.5-fold ratio in tendency to blame other people or external circumstances over own behavior or characteristics (P. Gilbert, J. Gilbert, & Irons, 2004).

As a section summary, the present modeling study, together with the reviewed previous research reports, suggest that depression offers a nonparticipation strategy to joint enterprises that are otherwise difficult to avoid, and that past evolutionary selection among groups of individuals, bands, or tribes, has favored this behavioral mode. This "nonparticipation hypothesis" serves to explain several seemingly unrelated empirical observations, such as association between income inequality and depression across populations, co-morbidity and proximal biological mechanisms for negative emotions, mechanisms for the social network-dynamics of emotion, existence of genes that 'interact' with environment in order to produce depression vulnerability thereby ensuring the potential for a nonparticipation response to defection in joint enterprises, and non-specific benefits of therapeutic alliance in depression treatment.

5.4 On the connections among the levels of analysis

According to the present results, and according to the other above-discussed research, individual depressive symptoms contain causally relevant information. Thus, depression researchers should therefore strive towards statistical descriptions of symptoms as opposed to descriptions of latent variables or diagnostic classes. Yet, depressive symptoms undeniably form a "positive manifold" [*i.e.*, are positively intercorrelated in populations (van der Maas et al., 2006)], and when sufficient theoretical prior information is lacking, statistical power issues often preclude complex models for individual symptoms in practice (section 3.4.3, 1st paragraph; Pashler & Wagenmakers, 2012; Simmons, Nelson, & Simonsohn, 2011; Wasserman, 2006). Hence, also realistic stochastic models for aggregates of symptoms (such as simple sums) are useful; models that acknowledge multiple background causes (Cramer et al., 2012; Hyde et al., 2008; Kendler et al., 2006) instead of simply assuming single latent constructs (Borsboom et

al., 2003; Molenaar & Campbell, 2009; Reise & Waller, 2009). The situation is not totally unlike that in statistical mechanics (*i.e.*, in physics), where one does not abandon the macroscopic concept of temperature after understanding that it represents only microscopic movements/energies of multiple particles, but brings this understanding into the models of temperature and studies the interfaces among the temperature, particles, and other macroscopic properties of physical systems (Baxter, 2007; Einstein, 1956).

Are there plausible relationships arising from the present studies on individual symptoms, their sum scores, and evolutionary models of depression? For example, the present theoretical model suggested that abstaining from joint enterprises (*e.g.*, work life) is causally linked with depression. A large empirical study found that unemployed men have more *stable* decreased life satisfaction than unemployed women (Clark et al., 2008), possibly reflecting larger selection strength of this enterprise among men than among women. On the stochastic aggregate-level, the present study found that men had more stable depressive symptoms than women, as would be expected. Edward Hagen's bargaining model motivating the nonparticipation hypothesis was originally developed in the context of post-partum depression (Hagen, 1999), and only then extended to general depression (Hagen, 2003). Hence, childbirth provides at least one obvious temporary risk factor for depression that is specific to women, compatible with the nonparticipation hypothesis, and contributes to the observed higher *volatility* in symptoms compared to men (Figure 3A). As another example, re-interpretation of the social pain hypothesis (MacDonald & Leary, 2005) in the context of the finding that body-image dissatisfaction is associated with chronic dysphoria (see section 5.2) can be made.

On one hand, this study found that dissatisfaction to one's appearances is associated with chronic dysphoria, equally strongly in both men and women. On the other hand, the theoretical part of this thesis suggested that unequal gains from equal efforts is associated with dysphoria and withdrawal from the enterprise (the nonparticipation hypothesis of depression). A large longitudinal population study of 11 407 participants has reported that physical appearance has a substantial effect on earnings and

employment patterns, mainly through employer discrimination (Harper, 2000). Furthermore, the effect was asymmetric so that adverse appearances decreased earnings more than attractiveness increased them; and contrary to popular belief, physical appearance was as important for men as it was for women (Harper, 2000). Hence, from the viewpoint of a less attractive person, there may often be realistic grounds to perceive defection in public goods; a perception that promotes depression, dysphoria, and nonparticipation according to the present theory. Therefore the present theory is consistent with the present empirical observation about chronicity and body-image dissatisfaction. Whereas appearance issues did not show gender differences in relation to chronic dysphoria, they did show a gender difference on their average level, as is commonly observed (Ålgars et al., 2009).

Sleep problems and tiredness were recognized as short-term causes for general dysphoria. Also this is in line with the nonparticipation hypothesis, as a tired person has less available energetic resources than others, and therefore a temporally elevated relative costs from taking part to joint enterprises of the society. The finding is also in line with women's higher variance and average in dysphoria, as women have 1.41-fold odds for insomnia compared to men (B. Zhang & Wing, 2006). Despite some temporally less-stable symptoms, in general, the individual depressive symptoms were surprisingly stable compared to their sum score, considering that low stability/reliability is a common justification for the use of latent-trait measurement models. When a several item measurement model is even approximately correct, it should considerably improve the reliability and stability over its constituent items. The fact that this did not happen for the symptom sum undermines the idea of latent-trait measurement. The temporal movements of sum score must reflect contributions from several sources (as in Brownian motion, for example). In general, combining and comparing research on many levels of analysis will help researchers to find robust theories and unifying principles. The above connections among the levels of analysis were intended only as examples rather than as an exhaustive listing.

What exactly is the benefit of multiple levels of analysis, with a fare amount of incompleteness in the overall picture? While not contradicting the network view, the

evolutionary viewpoint promises hypotheses based on known history. This may turn out critical, as the previously tested latent-cause models do not seem to fit to the data (section 1.1), but the purely statistical learning of multidimensional, possibly non-linear and certainly non-closed, dynamic network systems with unspecified connection strengths between only partially known nodes is a remarkably difficult task. That is, fitting a wrong model is misleading, and allowing too many degrees of freedom precludes efficient empirical assessment (*c.f.*, section 3.4.3, 1st paragraph), but the making of an 'educated guess' for a more correct model than the previous ones could allow progression; hence the provision of the nonparticipation hypothesis for the scrutiny of the scientific community.

5.5 Limitations

A general limitation in the empirical part of this thesis was its sole reliance on self-reported symptom scores. Depressive symptoms and body-image dissatisfaction are quite subjective, and simply asking about them is the gold standard of recording, although the methods of asking vary (see Introduction). For sleep problems, however, polysomnographic recordings could offer comprehensive recording of biophysical changes during sleep. Such cumbersome measures are rarely applied in large epidemiological studies, and unfortunately this study did not make an exception. An external opinion on physical appearance was also lacking, and this would have aided in differentiating between self-defeating cognitive biases and more realistic interpretations of appearances. This difference has important theoretical implications, and being able to discern the two alternatives would definitely be a strength for future studies.

As another potential limitation, the present study did not assess "body dysmorphic disorder" (Phillips & Crino, 2001). Most of the patients with body dysmorphic disorder also suffer from major depressive disorder (75% life-time and 61% current comorbodity), and many have social phobia [37% life-time and 31% current comorbidity (Gunstad & Phillips, 2003)]. As a specific psychiatric diagnosis, however,

body dysmorphic disorder is unlikely to account for the present findings, as its estimated prevalence in general populations is approximately 2% (Buhlmann et al., 2010; Koran, Abujaoude, Large, & Serpe, 2008). Even the most severe category of the "body-image dissatisfaction" item was selected by 5.3% of the sample herein, and only 26.7% reported never experiencing appearance concerns. A continuous tendency similar to the diagnostic entity may exist, however. The defining characteristic of body dysmorphic disorder, the preoccupation with imagined or slight defect in appearance, certainly leads to high scores in the present body-image dissatisfaction item. Furthermore, a 'defect in appearance' is a subjective and culture- and age-bound issue (Ålgars et al., 2009). As there are different aspects in body-image dissatisfaction (Ålgars et al., 2009), future studies might improve the present result by separately assessing these different aspects.

As a technical limitation, EDM parameter estimates were fairly inaccurate. More measurement points might facilitate better precision for the parameter estimates than what was obtained here, and might also support estimation of additional individual-level parameters (*e.g.*, instead the population-parameter A) along the lines of the DynAffect model (Kuppens et al., 2010). Less uncertainty in parameters would directly translate to less uncertainty in the model-derived variance decomposition. Unfortunately, present estimates had a high variance. The bootstrap estimates can be considered conservative, however, in the sense that they incorporate both the sampling variability and the variability due to algorithmic convergence issues; OpenMx cannot compute the usual sampling-theory estimates when nonlinear constraints required by the stochastic differential equation model are used in estimation (Boker et al., 2011; Oud & Delsing, 2010). In addition to variability issues, information was not available regarding who (if any) of the participants had received treatment for depression. The data represent the general population of Finland, however; with the exception that some of the more severe cases of depression are lost to study attrition (Table 2, and the related section of Results). Regarding another technical limitation, the simple method of Table 2 detected attrition effects on depressive symptoms that had magnitude $d > 0.2$. In contrast, the FIML method corrected only about a tenth of that. This suggests that FIML estimates

based on historical depressive symptoms are not informative enough to correct the majority of bias due to missing data in epidemiological depression research.

On the more conceptual side, this thesis reviewed empirical support for the nonparticipation hypothesis, but no claim was made for exclusiveness. Other pathways to depression are also likely to exist. In fact, depressed mood is the fourth most connected symptom in a network formed from psychiatric symptoms and diagnoses (Borsboom et al., 2011), and also some underlying genetic vulnerabilities appear shared with several other disorders (Cross-disorder Group of the Psychiatric Genomics Consortium, 2013a, 2013b). The lack of exclusive nature may render the nonparticipation hypothesis difficult to refute, whereas most useful models typically are those that yield clear predictions that are easy to confirm or refute. The provided general mathematical framework hopefully aids researchers in construction of quantitative predictions that can be refuted or confirmed, however.

In relation to its motivating Bargaining Model, the suggested nonparticipation hypothesis builds on the more feasible concept of behavioral imitation compared to the concept of (rational) bargaining strategy. The nonparticipation hypothesis puts depression in to a larger context of social dynamics, tapping to a wider range of empirical data, and it does not require as clear specification of a target of influence as bargaining. In presence of a large burden of adversities, targeted behavior could be a problematic interpretation. In contrast, the original bargaining interpretation with long-term fitness implications perhaps makes repeated suicide attempts easier to understand, although suicide attempt could feasibly signal nonparticipation as well. The precise relationships among nonparticipation, punishing, and bargaining remain to be further demarcated, however, both mathematically and conceptually.

5.6 Some practical implications

First, depressive symptoms should be treated within some other stochastic model than the commonly applied linear, unidimensional, latent-variable model. Their aggregates

may be usefully described by stochastic differential equation models, such as the presently applied EDM or the previously used DynAffect (Kuppens et al., 2010). EDM may be more fitting for epidemiologic panel-data sets with scarce temporal sampling, whereas the DynAffect provides more information from densely sampled time-series data. Both models can incorporate many established empirical observations into a single abstract representation. Such representations are important because it is too cumbersome for humans to mentally track all relevant baseline information in complex issues (Kahneman, 2003). Gradually incorporating more empirical observations to an increasingly refined mathematical model is a time-honored way to circumvent our cognitive limitations. Also, epidemiological information from models like EDM and DynAffect may aid in the construction of individual-level predictive models for psychotherapy purposes, as therapy-outcome prediction via stochastic differential equation models has already been suggested (Molenaar, 2010).

Second, it has been suggested that researchers should concentrate more on the dynamics of individual depressive symptoms (Cramer et al., 2010, 2012). This appears wise also in the light of the present findings, as body-image dissatisfaction was recognized as a useful predictor of chronic depression and/or dysphoria, and sleep problems as a potential short-term push toward generalized dysphoria. Currently it seems that depression in general, and body-image dissatisfaction and sleep problems specifically, can all be altered by non-invasive therapeutic treatments (Bower et al., 2013; Britton, Haynes, Fridel, & Bootzin, 2010; Cuijpers, Donker, van Straten, Li, & Andersson, 2010; Lynch et al., 2010; Wade, George, & Atkinson, 2009). Instead, somatic antidepressent treatments based-on latent cause idea “do not appear to be useful for our current MDD definitions” (Ghaemi et al., 2013), as they are low or negligible in effectiveness (Ghaemi et al., 2013; Kirsch et al., 2008; Pigott, Leventhal, Alter, & Boren, 2010) and may incur increases in several substantial health risks, including mortality risk (Andrews et al., 2012; Coupland et al., 2011; Ghaemi et al., 2013).

Third, social dynamics and subjectively unfair social conditions deserve to be studied more in the context of depression, according to the nonparticipation hypothesis and related accounts. Better understanding of these factors may allow us to recognize those

individuals that are likely to remit spontaneously (Whiteford et al., 2012) and those that may remit under social and/or financial guidance or help, thereby discriminating these inherently "adaptive" cases from true biological dysfynctions. Specific effects of all existing therapies for depression appear modest (Cuijpers, van Straten, Bohlmeijer, Hollon, & Andersson, 2010), whereas the societal significance of depression is large (Vos et al., 2013); therefore, novel research openings are sorely needed. Although it is premature to offer treatments based on the present study, the nonparticipation hypothesis suggests that by promoting social justice one might also decrease the prevalence of depression. Such attempts would concur with the recently promoted public health approach to prevention of depression (Ghaemi et al., 2013). In addition, the hypothesis implies that purely chemical extinguishing of all depressive symptoms from a population (Kleinman, 2012) may be a formidable task due to suggested evolutionary underpinnings of the condition; if achieved, it might bring about some undesired cooperation-evaporating side-effects for the wider society. For more than twenty years, some researchers have wondered "whether widespread use of antidepressants might be … tampering with the mechanisms that regulate human social hierarchies" (Nesse, 1991).

5.7 Conclusion

Depressive disorders are a badly-understood major public-health problem. This thesis aimed for better understanding by constructing a more empirically justified model for the progression of widely studied sum scores, by studying the dynamics of individual depressive symptoms, and by drawing from sociobiological models for evolution of cooperation. The first and second kinds of studies were based on data from several thousand participants on epidemiological studies in Finland and in the United States, whereas the latter kind of study was based on mathematical and literature analysis.

A mean-reverting stochastic differential equation model was estimated for the 16-year longitudinal population trajectories of depressive symptoms; body-image

dissatisfaction was recognized as risk factor for chronic dysphoria/depression, and sleep problems for more transient dysphoria; and the depressive symptoms were interpreted as providing a novel, biologically ingrained, route for abstaining from otherwise compulsory social and economic joint enterprises, whereas a strategy for such nonparticipation is shown to promote the emergence of cooperation among self-interested agents.

The theoretical and empirical findings of this thesis were not contradictory, but mostly formed mutually supportive layers of modeling and observing population dynamics of the depressive symptoms. Currently, large quantities of data exist in depression research, but there is a shortage of commonly accepted and unifying theory. One can only hope that the developments of this thesis will aid in building some sorely needed theory, with appropriate connections to empirical reality.

6 REFERENCES

Abdi, H. (2007). Bonferroni test. In *Salkind, N. J. (Ed.), Encyclopedia of Measurement and Statistics: Volume 1* (pp. 103–106). Thousand Oaks, California: SAGE Publications.

Almeida, O. P., Alfonso, H., Yeap, B. B., Hankey, G., & Flicker, L. (2011). Complaints of difficulty to fall asleep increase the risk of depression in later life: the Health in Men study. *Journal of Affective Disorders*, *134*, 208–216.

American Psychiatric Association. (2000). *DSM-IV-TR: Diagnostic and Statistical Manual of Mental Disorders. Text revision.* (4th ed.). Washington, USA: American Psychiatric Association.

Andrews, P. W., Thomson, J. A., Amstadter, A., & Neale, M. C. (2012). Primum non nocere: an evolutionary analysis of whether antidepressants do more harm than good. *Frontiers in Psychology*, *3*, 117.

Andrews, P. W., & Thomson Jr, J. A. (2009). The bright side of being blue: depression as an adaptation for analyzing complex problems. *Psychological Review*, *116*, 620–654.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, *211*, 1390–1396.

Bach, F. R., & Jordan, M. I. (2003). Kernel independent component analysis. *The Journal of Machine Learning Research*, *3*, 1–48.

Bartlett, M. S. (1946). On the theoretical specification and sampling properties of autocorrelated time-series. *Supplement to the Journal of the Royal Statistical Society*, *8*, 27–41.

Baxter, R. J. (2007). *Exactly Solved Models in Statistical Mechanics*. Mineola, New York: Dover Publications.

Beck, A. T. (1967). *Depression: Clinical, Experimental, and Theoretical Aspects*. Philadelphia, Pennsylvania: University of Pennsylvania Press.

Beck, A. T. (2008). The evolution of the cognitive model of depression and its neurobiological correlates. *American Journal of Psychiatry*, *165*, 969–977.

Beck, A. T., & Steer, R. A. (1993). *Manual for the Beck Depression Inventory.* San Antonio, Texas: Psychological Corporation.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, Texas: Psychological Corporation.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, *4*, 561–571.

Benden, A., & Farvolden, P. (2010). Depression and the workplace: a progress report. *Current Psychiatry Reports*, *101*, 73–79.

Bienvenu, O. J., Davydow, D. S., & Kendler, K. S. (2011). Psychiatric “diseases” versus behavioral disorders and degree of genetic influence. *Psychological Medicine*, *41*, 33–40.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D. U. (2006). Complex networks: Structure and dynamics. *Physics Reports*, *424*, 175–308.

Boker, S., Neale, M., Maes, H., Wilde, M., Spiegel, M., Brick, T., … Bates, T. (2011). OpenMx: an open source extended structural equation modeling framework. *Psychometrika*, *76*, 306–317.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York, USA: John Wiley & Sons, Inc.

Borsboom, D., Cramer, A. O. J., Schmittmann, V. D., Epskamp, S., & Waldorp, L. J. (2011). The small world of psychopathology. *PLoS One*, *6*, e27407.

Borsboom, D., Mellenbergh, G. J., & Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110*, 203–219.

Boswell, P. C., & Murray, E. J. (1981). Depression, schizophrenia, and social attraction. *Journal of Consulting and Clinical Psychology*, *49*, 641–647.

Bower, P., Kontopantelis, E., Sutton, A., Kendrick, T., Richards, D. A., Gilbody, S., … Liu, E. T.-H. (2013). Influence of initial severity of depression on effectiveness of low intensity interventions: meta-analysis of individual patient data. *British Medical Journal*, *346*, f540.

Bowles, S. (2009). Did warfare among ancestral hunter-gatherers affect the evolution of human social behaviors? *Science*, *324*, 1293–1298.

Britton, W. B., Haynes, P. L., Fridel, K. W., & Bootzin, R. R. (2010). Polysomnographic and subjective profiles of sleep continuity before and after mindfulness-based cognitive therapy in partially remitted depression. *Psychosomatic Medicine*, *72*, 539–548.

Bromet, E., Andrade, L. H., Hwang, I., Sampson, N. A., Alonso, J., de Girolamo, G., … Kessler, R. C. (2011). Cross-national epidemiology of DSM-IV major depressive episode. *BMC Medicine*, *9*, 90.

Brosnan, S. F., & De Waal, F. B. M. (2003). Monkeys reject unequal pay. *Nature*, *425*, 297–299.

Buhlmann, U., Glaesmer, H., Mewes, R., Fama, J. M., Wilhelm, S., Brähler, E., & Rief, W. (2010). Updates on the prevalence of body dysmorphic disorder: a population-based survey. *Psychiatry Research*, *178*, 171–175.

Buss, D. M. (2008). Human nature and individual differences: evolution of human personality. In: *O. P. John, R. W. Robins, & L. A. Pervin (Eds.) Handbook of Personality: Theory and Research*, (3rd ed.). New York, USA: The Guildford Press.

Carver, C. S., & Miller, C. J. (2006). Relations of serotonin function to personality: current views and a key methodological issue. *Psychiatry Research*, *144*, 1–15.

Caspi, A., Hariri, A. R., Holmes, A., Uher, R., & Moffitt, T. E. (2010). Genetic sensitivity to the environment: the case of the serotonin transporter gene and its implications for studying complex diseases and traits. *American Journal of Psychiatry*, *167*, 509–527.

Cesarini, D., Dawes, C. T., Johannesson, M., Lichenstein, P., & Wallace, B. (2009). Experimental game theory and behavior genetics. *Annals of the New York Academy of Sciences, 1167*, 66–75.

Chen, P., & Shakhnovich, E. (2010). Thermal adaptation of viruses and bacteria. *Biophysical Journal*, *98*, 1109–1118.

Cicchetti, D., & Rogosh, F. A. (1996). Equifinality and multifinality in developmental psychopathology. *Development and Psychopathology*, *8*, 597–600.

Clark, A. E., Diener, E., Georgellis, Y., & Lucas, R. E. (2008). Lags and leads in life satisfaction: a test of the baseline hypothesis. *The Economic Journal*, *118*, F222–F243.

Clarke, H., Flint, J., Attwood, A. S., & Munafò, M. R. (2010). Association of the 5-HTTLPR genotype and unipolar depression: a meta-analysis. *Psychological Medicine*, *40*, 1767–1778.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). New Jersey, USA: Lawrence Erlbaum Associates.

Cole, D. A., & Martin, N. C. (2005). The longitudinal structure of the Children's Depression Inventory: testing a latent trait-state model. *Psychological Assessment*, *17*, 144–155.

Coupland, C., Dhiman, P., Morriss, R., Arthur, A., Barton, G., & Hippisley-Cox, J. (2011). Antidepressant use and risk of adverse outcomes in older people: population based cohort study. *British Medical Journal*, *343*, d4551.

Coyne, J. C. (1976). Depression and the response of others. *Journal of Abnormal Psychology*, *85*, 186–193.

Cramer, A. O. J., Borsboom, D., Aggen, S. H., & Kendler, K. S. (2012). The pathoplasticity of dysphoric episodes: differential impact of stressful life events on the pattern of depressive symptom inter-correlations. *Psychological Medicine*, *42*, 957–965.

Cramer, A. O. J., Waldorp, L. J., van der Maas, H. L. J., & Borsboom, D. (2010). Comorbidity: a network perspective. *Behavioral and Brain Sciences*, *33*, 137–193.

Cross-disorder Group of the Psychiatric Genomics Consortium. (2013a). Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nature Genetics, Advanced Online Publication*, doi:10.1038/ng.2711

Cross-disorder Group of the Psychiatric Genomics Consortium. (2013b). Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *The Lancet*, *381*, 1371–1379.

Cuijpers, P., Donker, T., van Straten, A., Li, J., & Andersson, G. (2010). Is guided self-help as effective as face-to-face psychotherapy for depression and anxiety disorders? A systematic review and meta-analysis of comparative outcome studies. *Psychological Medicine*, *40*, 1943–1957.

Cuijpers, P., van Straten, A., Bohlmeijer, E., Hollon, S. D., & Andersson, G. (2010). The effects of psychotherapy for adult depression are overestimated: a meta-analysis of study quality and effect size. *Psychological Medicine*, *40*, 211–223.

Darwin, C. (1871). *The Descent of Man, and Selection in Relation to Sex*. London: John Murray.

Davies, S., Naik, P. C., & Lee, A. S. (2001). Depression, suicide, and the national service framework. *British Medical Journal*, *322*, 1500–1501.

De Jonge, P. (2012). Deconstructing depression: disentanglement of the phenomenology and etiology of a heterogeneous, multifactorial disorder. Project #OND1349284 in the national academic research

and collaborations information system of Netherlands. Retrieved in 1st of June 2013 from http://www.narcis.nl/research/RecordID/OND1349284/Language/en

De Jonge, P., & Roest, A. M. (2012). Depression and cardiovascular disease: the end of simple models. *The British Journal of Psychiatry*, *201*, 337–338.

De Silva, H., Hauert, C., Traulsen, A., & Sigmund, K. (2010). Freedom, enforcement, and the social dilemma of strong altruism. *Journal of Evolutionary Economics*, *20*, 203–217.

Disner, S. G., Beevers, C. G., Haigh, E. A. P., & Beck, A. T. (2011). Neural mechanisms of the cognitive model of depression. *Nature Reviews Neuroscience*, *12*, 467–477.

Efron, B., & Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. USA: Chapman & Hall/CRC.

Einstein, A. (1956). *Investigations on the Theory of the Brownian Movement* (Reprint of the first english translation from 1926.). USA: Dover, Inc.

Ekström, J. (2009). An empirical polychoric correlation. In *Contributions to the Theory of Measures of Association for Ordinal Variables (PhD Thesis)*. Retrieved from http://escholarship.org/uc/item/2tw0b313 (last access 17.4.2012)

Enders, C. K. (2001). The impact of nonnormality on full information maximum-likelihood estimation for structural equation models with missing data. *Psychological Methods*, *6*, 352–370.

Farmer, A., Mahmood, A., Redman, K., Harris, T., Sadler, S., & McGuffin, P. (2003). A sib-pair study of the Temperament and Character Inventory scales in major depression. *Archives of General Psychiatry*, *60*, 490–496.

Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *The American Economic Review*, *90*, 980–994.

Fiscella, K., & Franks, P. (2000). Individual income, income inequality, health, and mortality: what are the relationships? *Health Services Research*, *35*, 307–318.

Flückiger, C., Del Re, A. C., Wampold, B. E., Symonds, D., & Horvath, A. (2012). How central is the alliance in psychotherapy? A multilevel longitudinal meta-analysis. *Journal of Counseling Psychology, 59*, 10–17.

Ford, D. E., & Kamerow, D. B. (1989). Epidemiologic study of sleep disturbances and psychiatric disorders. *The Journal of the American Medical Association*, *262*, 1479–1484.

Fowler, J. H. (2005). Altruistic punishment and the origin of cooperation. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 7047–7049.

Fowler, J. H., & Christakis, N. A. (2008). The dynamic spread of happiness in a large social network. *BMJ: British medical journal*, *337*, a2338.

Fowler, J. H., & Christakis, N. A. (2010). Cooperative behavior cascades in human social networks. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 5334–5338.

Friedman, R. A. (2012, January 16). Depression defies rush to find evolutionary upside. *The New York Times*. Retrieved from http://www.nytimes.com/2012/01/17/health/depression-defies-rush-to-find-evolutionary-upside.html

Galambos, N. L., Barker, E. T., & Krahn, H. J. (2006). Depression, self-esteem, and anger in emerging adulthood: seven-year trajectories. *Developmental Psychology*, *42*, 350–365.

Gelman, A., & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York: Cambridge University Press.

Genest, C., & Favre, A. (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering*, *12*, 347–368.

Ghaemi, S. N., Vöhringer, P. A., & Whitham, E. A. (2013). Antidepressants from a public health perspective: re-examining effectiveness, suicide, and carcinogenicity. *Acta Psychiatrica Scandinavica*, *127*, 89–93.

Gilbert, P. (2006). Evolution and depression: issues and implications. *Psychological Medicine*, *36*, 287–297.

Gilbert, P., Gilbert, J., Irons, C. (2004). Life events, entrapments and arrested anger in depression. *Journal of Affective Disorders, 79*, 149–160.

Godin, I., Kittel, F., Coppieters, Y., & Siegrist, J. (2005). A prospective study of cumulative job stress in relation to mental health. *BMC Public Health*, *5*, 67.

Godoy, R. A., Reyes-García, V., McDade, T., Huanca, T., Leonard, W. R., Tanner, S., & Vadez, V. (2006). Does village inequality in modern income harm the psyche? Anger, fear, sadness, and alcohol consumption in a pre-industrial society. *Social Science & Medicine*, *63*, 359–372.

Grabe, S., Ward, L. M., & Hyde, J. S. (2008). The role of the media in body image concerns among women: a meta-analysis of experimental and correlational studies. *Psychological Bulletin*, *134*, 460–476.

Gracia-Lázaro, C., Ferrer, A., Ruiz, G., Tarancón, A., Cuesta, J. A., Sánchez, A., & Moreno, Y. (2012). Heterogeneous networks do not promote cooperation when humans play a Prisoner's Dilemma. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 12922–12926.

Gunstad, J., & Phillips, K. A. (2003). Axis I comorbidity in body dysmorphic disorder. *Comprehensive Psychiatry*, *44*, 270–276.

Gächter, S., Renner, E., & Sefton, M. (2008). The long-run benefits of punishment. *Science*, *322*, 1510.

Hagen, E. H. (1999). The functions of postpartum depression. *Evolution and Human Behavior*, *20*, 325–359.

Hagen, E. H. (2003). The bargaining model of depression. In *Hammerstein, P. (Ed.), Genetic and Cultural Evolution of Cooperation* (pp. 97–123). Cambridge, Massachusetts, USA: The MIT Press.

Hagen, E. H. (2011). Evolutionary theories of depression: a critical review. *Canadian Journal of Psychiatry*, *56*, 716–726.

Hammerstein, P., Hagen, E. H., Herz, A. W. M., & Herzel, H. (2006). Robustness: a key to evolutionary design. *Biological Theory, 1*, 90–93.

Hankin, B. L., Abramson, L. Y., Moffitt, T. E., Silva, P. A., McGee, R., & Angell, K. E. (1998). Development of depression from preadolescence to young adulthood: emerging gender differences in a 10-year longitudinal study. *Journal of Abnormal Psychology*, *107*, 128–140.

Hansenne, M., Reggers, J., Pinto, E., Kjiri, K., Ajamier, A., & Ansseau, M. (1999). Temperament and character inventory (TCI) and depression. *Journal of Psychiatric Research*, *33*, 31–36.

Harper, B. (2000). Beauty, stature and the labour market: a British cohort study. *Oxford Bulletin of Economics and Statistics*, *62*, 771–800.

Haslam, N., Holland, E., & Kuppens, P. (2012). Categories versus dimensions in personality and psychopathology: a quantitative review of taxometric research. *Psychological Medicine*, *42*, 903–920.

Hauert, C., Traulsen, A., Brandt, H., Nowak, M. A., & Sigmund, K. (2007). Via freedom to coercion: the emergence of costly punishment. *Science*, *316*, 1905–1907.

Hauser, M. R., Sewell, W. H., & Hauser, T. S. (2006). Wisconsin Longitudinal Study (WLS) [graduates]: 1992/93. [machinereadable data file]. Madison, WI: University of Wisconsin-Madison, WLS. [distributor]. Retrieved from http://www.ssc.wisc.edu/~wls/documentation/

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., … Henrich, N. (2006). Costly punishment across human societies. *Science*, *312*, 1767–1771.

Higham, D. J. (2001). An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Review*, *43*, 525–546.

Hill, A. B. (1965). The environment and disease: association or causation? *Proceedings of the Royal Society of Medicine*, *58*, 295–300.

Hill, A. L., Rand, D. G., Nowak, M. A., & Christakis, N. A. (2010). Emotions as infectious diseases in a large social network: the SISa model. *Proceedings of the Royal Society B: Biological Sciences*, *277*, 3827–3835.

Hyde, J. S., Mezulis, A. H., & Abramson, L. Y. (2008). The ABCs of depression: integrating affective, biological, and cognitive models to explain the emergence of the gender difference in depression. *Psychological Review*, *115*, 291–313.

Hyvärinen, A. (2010). Pairwise measures of causal direction in linear non-Gaussian acyclic models. In *JMLR Conference and Workshop Proceedings* (Vol. 13, pp. 1–16). Retrieved from (Last accessed: 1.6.2013) http://www.cs.helsinki.fi/u/ahyvarin/papers/ACML2010.pdf

Hyvärinen, A., Karhunen, J., & Oja, E. (2001). *Independent Component Analysis.* Wiley, New York.

Hyvärinen, A., & Smith, S. M. (2013). Pairwise likelihood ratios for estimation of non-Gaussian structural equation models. *Journal of Machine Learning Research*, *14*, 111−152.

Jansson-Fröjmark, M., & Lindblom, K. (2008). A bidirectional relationship between anxiety and depression, and insomnia? A prospective study in the general population. *Journal of Psychosomatic Research*, *64*, 443–449.

Jokela, M., Singh-Manoux, A., Shipley, M. J., Ferrie, J. E., Gimeno, D., Akbaraly, T. N., … Kivimäki, M. (2011). Natural course of recurrent psychological distress in adulthood. *Journal of Affective Disorders*, *130*, 454–461.

Jylhä, P., Mantere, O., Melartin, T., Suominen, K., Vuorilehto, M., Arvilommi, P., … Isometsä, E. (2011). Differences in temperament and character dimensions in patients with bipolar I or II or major depressive disorder and general population subjects. *Psychological Medicine*, *41*, 1579–91.

Kahneman, D. (2003). A perspective on judgement and choice: mapping bounded rationality. *American Psychologist*, *58*, 697–720.

Karg, K., Burmeister, M., Shedden, K., & Sen, S. (2011). The serotonin transporter promoter variant (5-HTTLPR), stress, and depression meta-analysis revisited: evidence of genetic moderation. *Archives of General Psychiatry*, *68*, 444–454.

Karpansalo, M., Kauhanen, J., Lakka, T. A., Manninen, P., Kaplan, G. A., & Salonen, J. T. (2005). Depression and early retirement: prospective population based study in middle aged men. *Journal of Epidemiology and Community Health*, *59*, 70–74.

Karsten, J., Hartman, C. A., Ormel, J., Nolen, W. A., & Penninx, B. W. J. H. (2010). Subthreshold depression based on functional impairment better defined by symptom severity than by number of DSM-IV symptoms. *Journal of Affective Disorders*, *123*, 230–237.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773–795.

Katainen, S., Räikkönen, K., & Keltikangas-Järvinen, L. (1999). Adolescent temperament, perceived social support, and depressive tendencies as predictors of depressive tendencies in young adulthood. *European Journal of Personality*, *13*, 183–207.

Katainen, S., Räikkönen, K., Keskivaara, P., & Keltikangas-Järvinen, L. (1999). Maternal child-rearing attitudes and role satisfaction and children's temperament as antecedents of adolescent depressive tendencies: follow-up study of 6- to 15-year-olds. *Journal of Youth and Adolescence*, *28*, 139–163.

Keltikangas-Järvinen, L., & Jokela, M. (2010). Nature and nurture in personality. *Focus*, *8*, 180–186.

Kendler, K. S., Eaves, L. J., Loken, E. K., Pedersen, N. L., Middeldorp, C. M., Reynolds, C., Boomsma, D., Lichtenstein P., Silberg, J., Gardner, G. O. (2011). The impact of environmental experiences on symptoms of anxiety and depression across the life span. *Psychological Science, 22*, 1343–1352.

Kendler, K. S., & Gardner, C. O. (2010). Dependent stressful life events and prior depressive episodes in the prediction of major depression: the problem of causal inference in psychiatric epidemiology. *Archives of General Psychiatry*, *67*, 1120–1127.

Kendler, K. S., & Gardner, C. O. (2011). A longitudinal etiologic model for symptoms of anxiety and depression in women. *Psychological Medicine, 41*, 2035–2045.

Kendler, K. S., Gardner, C. O., & Prescott, C. A. (2006). Toward a comprehensive developmental model for major depression in men. *American Journal of Psychiatry*, *163*, 115–124.

Kendler, K. S., Karkowski, L. M., & Prescott, C. A. (1998). Stressful life events and major depression: risk period, long-term contextual threat, and diagnostic specificity. *The Journal of Nervous and Mental Disease*, *186*, 661–669.

Kennan, J. & Wilson, R. (1993). Bargaining with private information. *Journal of Economic Literature, 31*, 45–104.

Kenny, D. A., & Zautra, A. (1995). The trait-state-error model for multiwave data. *Journal of Consulting and Clinical Psychology*, *63*, 52–59.

Kirsch, I., Deacon, B. J., Huedo-Medina, T. B., Scoboria, A., Moore, T. J., & Johnson, B. T. (2008). Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Medicine*, *5*, e45.

Kivimäki, M., Vahtera, J., Kawachi, I., Ferrie, J. E., Oksanen, T., Joensuu, M., Pentti, J., Salo, P., Elovainio, M., & Virtanen, M. (2010). Psychosocial work environment as a risk factor for absence with a psychiatric diagnosis: an instrumental-variables analysis. *American Journal of Epidemiology, 172*, 167–172.

Kleinman, A. (2012). Culture, bereavement, and psychiatry. *The Lancet*, *379*, 608–609.

Klenke, A. (2008). *Probability Theory: A Comprehensive Course*. London, England: Springer-Verlag.

Koran, L. M., Abujaoude, E., Large, M. D., & Serpe, R. T. (2008). The prevalence of body dysmorphic disorder in the United States adult population. *CNS Spectrums*, *13*, 316–322.

Kuppens, P., Oravecz, Z., & Tuerlinckx, F. (2010). Feelings change: accounting for individual differences in the temporal dynamics of affect. *Journal of Personality and Social Psychology*, *99*, 1042–1060.

Lawley, D. N., & Maxwell, A. E. (1971). *Factor Analysis as a Statistical Method.* (2nd ed.). London: Butterworths & Co.

Lester, D., & Gatto, J. L. (1990). Interpersonal trust, depression, and suicidal ideation in teenagers. *Psychological Reports*, *67*, 786–786.

Lilliefors, H. W. (1967). On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, *62*, 399–402.

Lord, F. M., & Novick, M. R. (1968). *Statistical Theories of Mental Test Scores with Contributions by A. Birnbaum* (1st ed.). Reading, Massachusetts: Addison-Wesley Publishing Company.

Lynch, D., Laws, K. R., & McKenna, P. J. (2010). Cognitive behavioural therapy for major psychiatric disorder: does it really work? A meta-analytical review of well-controlled trials. *Psychological Medicine*, *40*, 9–24.

MacDonald, G., & Leary, M. R. (2005). Why does social exclusion hurt? The relationship between social and physical pain. *Psychological Bulletin*, *131*, 202–223.

Marijnissen, G., Tuinier, S., Sijben, A. E. S., & Verhoeven, W. M. A. (2002). The temperament and character inventory in major depression. *Journal of Affective Disorders*, *70*, 219–223.

Merikangas, K. R., Zhang, H., Avenevoli, S., Acharyya, S., Neuenschwander, M., & Angst, J. (2003). Longitudinal trajectories of depression and anxiety in a prospective community study: the Zurich Cohort Study. *Archives of General Psychiatry*, *60*, 993–000.

Markon, K. E., Chmielewski, M., & Miller, C. J. (2011). The reliability and validity of discrete and continuous measures of psychopathology: a quantitative review. *Psychological Bulletin*, *137*, 856–879.

Martin, D. J., Garske, J. P., & Davis, M. K. (2000). Relation of therapeutic alliance with outcome and other variables: a meta-analytic review. *Journal of Consulting and Clinical Psychology, 68*, 438–450.

Merjonen, P., Keltikangas-Järvinen, L., Jokela, M., Seppälä, I., Lyytikäinen, L. P., Pulkki-Råback, L., … Ripatti, S. (2011). Hostility in adolescents and adults: a genome-wide association study of the Young Finns. *Translational Psychiatry*, *1*, e11.

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: bringing the person back into scientific psychology, this time forever. *Measurement*, *2*, 201–218.

Molenaar, P. C. M. (2010). Note on optimization of individual psychotherapeutic processes. *Journal of Mathematical Psychology*, *54*, 208–213.

Molenaar, P. C. M., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychological Science*, *18*, 112–117.

Muramatsu, N. (2003). County-level income inequality and depression among older americans. *Health Services Research*, *38*, 1863–1884.

Muthén, B., Kaplan, D., & Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. *Psychometrika*, *52*, 431–462.

Mykletun, A., Bjerkeset, O., Dewey, M., Prince, M., Overland, S., & Stewart, R. (2007). Anxiety, depression, and cause-specific mortality: the HUNT study. *Psychosomatic Medicine*, *69*, 323–331.

Mykletun, A., Bjerkeset, O., Øverland, S., Prince, M., Dewey, M., & Stewart, R. (2009). Levels of anxiety and depression as predictors of mortality: the HUNT study. *The British Journal of Psychiatry*, *195*, 118–125.

Nelsen, R. B. (2006). *An Introduction to Copulas* (2nd ed.). New York, NY: Springer.

Nesse, R. M. (1991). What good is feeling bad? The evolutionary benefits of psychic pain. *Sciences*, *31*, 30–37.

Nesse, R. M. (2000). Is depression an adaptation? *Archives of General Psychiatry*, *57*, 14–20.

Netterstrøm, B., Conrad, N., Bech, P., Fink, P., Olsen, O., Rugulies, R., & Stansfeld, S. (2008). The relation between work-related psychosocial factors and the development of depression. *Epidemiologic Reviews*, *30*, 118–132.

Nettle, D. (2004). Evolutionary origins of depression: a review and reformulation. *Journal of Affective Disorders*, *81*, 91–102.

Nowak, M. A. (2006). *Evolutionary Dynamics: Exploring the Equations of Life*. Massachusetts, Cambridge: The Belknap Press of Harvard University Press.

Nowak, M. A. (2012). Evolving cooperation. *Journal of Theoretical Biology*, *299*, 1–8.

Oravecz, Z., Tuerlinckx, F., & Vandekerckhove, J. (2011). A hierarchical latent stochastic differential equation model for affective dynamics. *Psychological Methods*, *16*, 468 – 490.

Oud, J. H. L., & Delsing, M. J. M. H. (2010). Continuous time modeling of panel data by means of SEM. In K. Montfort, J. Oud, & A. Satorra (Eds.), *Longitudinal Research with Latent Variables* (pp. 201–244). Berlin, Germany: Springer-Verlag.

Oud, J. H. L., & Jansen, R. A. R. G. (2000). Continuous time state space modeling of panel data by means of SEM. *Psychometrika*, *65*, 199–215.

Oud, J. H. L., & Singer, H. (2008). Continuous time modeling of panel data: SEM versus filter techniques. *Statistica Neerlandica*, *62*, 4–28.

Pashler, H., & Wagenmakers, E. (2012). Editors' introduction to the special section on replicability in psychological science: a crisis of confidence? *Perspectives on Psychological Science*, *7*, 528–530.

Paunio, T., Korhonen, T., Hublin, C., Partinen, M., Kivimäki, M., Koskenvuo, M., & Kaprio, J. (2009). Longitudinal study on poor sleep and life dissatisfaction in a nationwide cohort of twins. *American Journal of Epidemiology*, *169*, 206–213.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge, UK: Cambridge University Press.

Pearson, K. (1900). Mathematical contributions to the theory of evolution. VII. On the correlation of characters not quantitatively measurable. *Philosophical Transactions of the Royal Society of London*, *195*, 1–47.

Perlis, M. L., Giles, D. E., Buysse, D. J., Tu, X., & Kupfer, D. J. (1997). Self-reported sleep disturbance as a prodromal symptom in recurrent depression. *Journal of Affective Disorders*, *42*, 209–212.

Phillips, K. A., & Crino, R. (2001). Body dysmorphic disorder. *Current Opinion in Psychiatry*, *14*, 113–118.

Pickett, K. E., James, O. W., & Wilkinson, R. G. (2006). Income inequality and the prevalence of mental illness: a preliminary international analysis. *Journal of Epidemiology and Community Health*, *60*, 646–647.

Pigott, H. E., Leventhal, A. M., Alter, G. S., & Boren, J. J. (2010). Efficacy and effectiveness of antidepressants: current status of research. *Psychotherapy and Psychosomatics*, *79*, 267–279.

Pirkola, S. P., Isometsä, E., Suvisaari, J., Aro, H., Joukamaa, M., Poikolainen, K., … Lönnqvist, J. K. (2005). DSM-IV mood-, anxiety- and alcohol use disorders and their comorbidity in the Finnish general population. *Social Psychiatry and Psychiatric Epidemiology*, *40*, 1–10.

Price, J. S., Gardner, R., & Erickson, M. (2004). Can depression, anxiety, and somatization be understood as appeasement displays? *Journal of Affective Disorders*, *79*, 1–11.

Puurtinen, M., & Mappes, T. (2009). Between-group competition and human cooperation. *Proceedings of the Royal Society B: Biological Sciences*, *276*, 355–360.

R Core Team. (2012). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Radloff, L. S. (1977). The CES-D scale: a self report depression scale for research in the general population. *Applied Psychological Measurement*, *1*, 385–401.

Raitakari, O. T., Juonala, M., Rönnemaa, T., Keltikangas-Järvinen, L., Räsänen, L., Pietikäinen, M., … Marniemi, J. (2008). Cohort profile: the Cardiovascular Risk in Young Finns study. *International Journal of Epidemiology*, *37*, 1220–1226.

Reise, S. P., & Waller, N. G. (2009). Item response theory and clinical measurement. *Annual Review of Clinical Psychology*, *5*, 27–48.

Rizopoulos, D. (2006). ltm: an R package for latent variable modeling and item response theory analyses. *Journal of Statistical Software*, *17*, 1–25.

Rosenquist, J. N., Fowler, J. H., & Christakis, N. A. (2010). Social network determinants of depression. *Molecular Psychiatry*, *16*, 273–281.

Samejima, F. (1997). Graded response model. In *van der Linden, W. J., Hambleton, R. K. (Eds.) Handbook of Modern Item Response Theory* (pp. 85–100). New York: Springer.

Shafer, A. B. (2006). Meta-analysis of the factor structures of four depression questionnaires: Beck, CES-D, Hamilton, and Zung. *Journal of Clinical Psychology*, *62*, 123–146.

Shimizu, S., Hoyer, P. O., Hyvärinen, A., & Kerminen, A. (2006). A linear non-Gaussian acyclic model for causal discovery. *Journal of Machine Learning Research*, *7*, 2003–2030.

Shimizu, S., Inazumi, T., Sogawa, Y., Hyvärinen, A., Kawahara, Y., Washio, T., … Bollen, K. (2011). DirectLiNGAM: a direct method for learning a linear non-Gaussian structural equation model. *Journal of Machine Learning Research*, *12*, 1225–1248.

Sigmund, K. (2010). *The Calculus of Selfishness*. Princeton, New Jersey: Princeton University Press.

Sigmund, K., De Silva, H., Traulsen, A., & Hauert, C. (2010). Social learning promotes institutions for governing the commons. *Nature*, *466*, 861–863.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*, 1359–1366.

Sivertsen, B., Salo, P., Mykletun, A., Hysing, M., Pallesen, S., Krokstad, S., … Øverland, S. (2012). The bidirectional association between depression and insomnia: the HUNT study. *Psychosomatic Medicine*, *74*, 758–765.

Sloman, L. (2008). A new comprehensive evolutionary model of depression and anxiety. *Journal of Affective Disorders*, *106*, 219–228.

Solomon, A., Haaga, D. A. F., & Arnow, B. A. (2001). Is clinical depression distinct from subthreshold depressive symptoms? A review of the continuity issue in depression research. *The Journal of Nervous and Mental Disease*, *189*, 498–506.

Sturm, R., & Gresenz, C. R. (2002). Relations of income inequality and family income to chronic medical conditions and mental health disorders: national survey. *British Medical Journal*, *324*, 1–5.

Sullivan, P. F., Neale, M. C., & Kendler, K. S. (2000). Genetic epidemiology of major depression: review and meta-analysis. *American Journal of Psychiatry*, *157*, 1552–1562.

Surbey, M. K. (2011). Adaptive significance of low levels of self-deception and cooperation in depression. *Evolution and Human Behavior*, *32*, 29–40.

Surbey, M. K., & McNally, J. J. (1997). Self-deception as a mediator of cooperation and defection in varying social contexts described in the iterated prisoner's dilemma. *Evolution and Human Behavior*, *18*, 417–435.

Surbey, M. K., & Simpson, J. (2010). Cheerful cheaters beware! *Presentation at The Human Behavior and Evolution Society Annual Meeting.* Abstrac retreived from: http://www.hbes.com/conference/pdf/conference_22.pdf. Last accessed: 11.1.2013.

Surtees, P. G., & Wainwright, N. W. J. (1999). Surviving adversity: event decay, vulnerability and the onset of anxiety and depressive disorder. *European Archives of Psychiatry and Clinical Neuroscience*, *249*, 86–95.

Sääksvuori, L., Mappes, T., & Puurtinen, M. (2011). Costly punishment prevails in intergroup conflict. *Proceedings of the Royal Society B: Biological Sciences*, *278*, 3428–3436.

Turchin, P., Currie, T. E., Turner, E. A. L., & Gavrilets, S. (2013). War, space, and the evolution of Old World complex societies. *Proceedings of the National Academy of Sciences of the United States of America, Early edition*, doi:10.1073/pnas.1308825110

Wade, T., George, W. M., & Atkinson, M. (2009). A randomized controlled trial of brief interventions for body dissatisfaction. *Journal of Consulting and Clinical Psychology*, *77*, 845–854.

Van de Waal, E., Borgeaud, C., & Whiten, A. (2013). Potent social learning and conformity shape a wild primate's foraging decisions. *Science*, *340*, 483–485.

Van der Maas, H. L. J., Dolan, C. V., Grasman, R. P. P. P., Wicherts, J. W., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: the positive manifold of intelligence by mutualism. *Psychological Review*, *113*, 842– 861.

Van Loo, H. M., de Jonge, P., Romeijn, J., Kessler, R. C., & Schoevers, R. A. (2012). Data-driven subtypes of major depressive disorder: a systematic review. *BMC Medicine*, *10*, 156.

Wasserman, L. (2006). *All of Nonparametric Statistics.* New York: Springer-Verlag.

Watson, P. J., & Andrews, P. W. (2002). Toward a revised evolutionary adaptationist analysis of depression: the social navigation hypothesis. *Journal of Affective Disorders*, *72*, 1–14.

Weich, S., Lewis, G., & Jenkins, S. P. (2001). Income inequality and the prevalence of common mental disorders in Britain. *British Medical Journal*, *178*, 222–227.

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: issues and guidance for practice. *Statistics in Medicine*, *30*, 377–399.

Whiteford, H. A., Harris, M. G., McKeon, G., Baxter, A., Pennell, C., Barendregt, J. J., & Wang, J. (2012). Estimating remission from untreated major depression: a systematic review and meta-analysis. *Psychological Medicine*, *E-publication Ahead of Print*, 1–17. doi:10.1017/S0033291712001717

Wiebe, S. T., Cassof, J., & Gruber, R. (2012). Sleep patterns and the risk for unipolar depression: a review. *Nature and Science of Sleep*, *4*, 63–71.

World Health Organization. (2010). WHO | International Classification of Diseases (ICD). *WHO*. Retrieved April 9, 2013, from http://www.who.int/classifications/icd/en/

Vos, T., Flaxman, A. D., Naghavi, M., Lozano, R., Michaud, C., Ezzati, M., … Aboyans, V. (2013). Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, *380*, 2163–2196.

Ylipaavalniemi, J., Kivimäki, M., Elovainio, M., Virtanen, M., Keltikangas-Järvinen, L., & Vahtera, J. (2005). Psychosocial work characteristics and incidence of newly diagnosed depression: a prospective cohort study of three different models. *Social Science & Medicine*, *61*, 111–122.

Zhang, B., & Wing, Y. (2006). Sex differences in insomnia: a meta-analysis. *Sleep*, *29*, 85–93.

Zhang, T., & Meaney, M. J. (2010). Epigenetics and the environmental regulation of the genome and its function. *Annual Review of Psychology*, *61*, 439–466.

Zimmerman, M., McGlinchey, J. B., Chelminski, I., Young, D. (2006). Diagnosing Major Depressive Disorder V: applying the DSM-IV exclusion criteria in clinical practice. *Journal of Nervous and Mental Disease, 194*, 530–533.

Åkerblom, H. K., Uhari, M., Pesonen, E., Dahl, M., Kaprio, E. A., Nuutinen, E. M., … Viikari, J. (1991). Cardiovascular risk in young Finns. *Annals of Medicine*, *23*, 35–39.

Ålgars, M., Santtila, P., Varjonen, M., Witting, K., Johansson, A., Jern, P., & Sandnabba, N. K. (2009). The adult body: how age, gender, and body mass index are related to body image. *Journal of Aging and Health*, *21*, 1112–1132.

Øksendal, B. K. (2003). *Stochastic Differential Equations: an Introduction with Applications*. Berlin, Heidelberg: Springer-Verlag.