Exploratory time series modeling of individuals with burn-out

Abstract

In psychology, psychopathology is usually conceptualized with latent variable models. However, a causal model of symptoms (where symptoms cause each other) seems a more plausible model to represent psychopathology. We wish to investigate if burn-out in an individual can be represented as such a model. Investigating the relationships between symptoms could provide valuable information on the workings of burnout in individuals. This was investigated using multivariate time series models. However, fitting the models proved not a trivial task; results found could not give ground to solid conclusions. Possible causes of the results are discussed, and based on the limitations of this study some directions are given for the future.

Internship Report: First Draft

name	:	Noémi Katalin Schuurman

Supervisor

within RMP (obligatory)	:	D. Borsboom
specialization	:	Methodology
research center / location	:	University of Amsterdam

Contents

Exploratory time series modeling of individuals with burn-out

In psychology, psychopathology is usually conceptualized with latent variable models. A latent variable is unobserved, and is usually measured indirectly through other observable variables. For instance, depression is measured by observed symptoms patients display. The basic idea is that the latent variable (depression) is a common cause of these observable variables (symptoms). This means that before the latent variable is known (added to the model) the observed variables will be correlated. When the latent variable is known it will explain the associations between the observed variables and the spurious correlations will disappear (local independence).

However, latent variable models may not plausibly represent psychopathology. For instance, it is unlikely that correlations between observed variables are spurious. It is likely that the depression symptoms 'sleep disturbances' and 'fatigue' will directly affect each other. So, the relationships between the observed variables can be explained by simple cause and effect, instead of a common latent variable.

In other words, causal models of symptoms may be another more plausible way to represent psychopathology than the latent variable model (Borsboom, 2008). Causal models of symptoms seem to reflect how practicing clinical psychologists view patient's disorders (Kim & Ahn, 2002). Also, it has been shown through simulation that even though a causal model of observable variables was the 'true' model, a latent variable model can fit just as well (van der Maas, Dolan, Grasman, Wicherts, Huizenga, & Raijmakers, 2006), so even though latent variable models fit psychopathology data, it does not mean it is in fact what is really happening. Furthermore, since the causal models do not need latent entities, it is the more simple model and should be preferred based on the Occam's Razor principle. However, even though causal models are conceptually more simple, they are not easily tested. This will require a new methodological and psychometric approach.

Time series models may offer one way to investigate the causal relationships between observed variables. In time series analyses, associations between variables can be modeled through time for an individual. This offers advantages over the usual multi-subject analyses utilized in psychology. Firstly, the variables are modeled through time, it is possible to demonstrate antecedence, a necessary condition for causality. Secondly, the analyses can be done for individuals. This seems sensible for research on psychopathology, since ultimately, clinical research is conducted in order to help individuals. Results from research employing multi-subject analyses do not necessarily generalize to any individual. By conducting the research on the level of the individual, this is no issue.

My task for the internship was to explore ways to represent psychopathology data as a causal model using time series modeling. In order to do this, an existing dataset was acquired from previous research on burnout.

Burn-out is a disorder characterized by severe and chronic fatigue, and cynical attitude towards work and work-related self-efficacy. The disorder is often accompanied by somatic symptoms or symptoms of depression (Sonnenschein, Sorbi, van Doornen, & Maas, 2006). Burn-out is not part of the DSM, since it is not yet internationally recognized as a valid disorder. However, in the Netherlands around 4% of the working population is estimated to have a severe burnout, while another 16% shows mild symptoms and is at risk of developing a severe burnout (Houtman, Schaufeli & Taris, 2000). A clinical burn-out can last for years, so it can have a tremendous effect on the patients life, as well as on the work society as a whole due to increases in sick leave (Kant, Jansen, Van Amelsvoort, Mohren & Swaen, 2004; Borritz, Rugulies, Christensen, Villadsen & Kristensen, 2006). Treatment has not been found to be very effective in diminishing burn-out symptoms (Blonk, Brenninkmeijer, Lagerveld & Houtman, 2006).

If more was known about the causal relationships between symptoms, perhaps treatment could be applied more effectively. Investigating the associations between burn-out symptoms using time series modeling could provide insight in the workings of burnout, and thus valuable information for clinical practice.

Methods

Participants and procedure

An existing dataset was acquired from the university of Utrecht. The data was originally part of the promotion research 'Sick with Burnout, clarified through electronic diaries' by Dr. M.A. Sonnenschein (2007). The sample consists of 60 participants with clinical burn-out and 40 healthy controls.

The burned-out participants were recruited through the Internet and burn-out treatment centers. Controls were recruited through newspaper advertisements and personal contacts. The participants were screened with the Maslach Burnout Inventory General Survey (MBI-GS; Schaufeli & Van Dierendonck, 2000), the Checklist Individual Strength (CIS; Bültmann, Vries , Beurskens, Bleijenberg, Vercoulen et al., 2000) and the Symptom Checklist-90-R (SCL-90-R; Arrindell & Ettema, 2002). Participants that showed a high level of exhaustion (MBI-GS exhaustion > 2.2), and a high level of cynicism (MBI-GS cynicism > 2) or a low level of personal accomplishment (MBI-GS < 3.67), and a high level of fatigue (CIS fatigue ≥ 76) were classified as having clinical burn-out. Healthy controls had to show no burn-out symptomatology; MBI-GS exhaustion <2.2), MBI-GS cynicism <2, MBI-GS personal accomplishment >3.67, CIS fatigue ≤ 76, SCL-90-R < 183. Subjects were excluded when they were diagnosed with other psychopathology than burnout (SCL-90-R), showed extremely severe psychopathology (SCL-90-

R), used antidepressants or anxiolytics, or if they were pregnant.

Selected participants signed an informed consent form. They received instructions for the electronic diary at home, and were approached two days later to asses if technical support was necessary. Technical support was also available during the following weeks of data collection. Participants filled in the questionnaires in the electronic diaries every morning after waking, every evening before bedtime, and randomly throughout the day, for two weeks. It concerned different questionnaires for the mornings, evenings and daytimes. On average, each participant filled in 13 evening and morning questionnaires, and 60 randomly administered questionnaires. Next to this the participants filled in a questionnaire concerning general information about the participants life situation once. Half a year later this process was repeated for the burned-out participants for another two weeks as a follow-up.

After the two weeks the participants were debriefed and received a remuneration of € 25. A more detailed description of the screening and data collection procedure can be found in the 2007 paper by Sonnenschein, Sorbi, van Doornen, Schaufeli, and Maas.

Materials

The questionnaire concerning general information about the participants life situation included questions about the participants sex, age, education-level, job, family situation (marital status, number of children) and health-related habits (exercise, smoking, drinking, medication).

The morning questionnaire contained questions about how the participant slept, and how they felt about the coming day. The evening questionnaire contained questions about the participants feelings about their accomplishments that day, how busy their day was and how many rest they needed. The randomly administered questionnaire contained questions about what they were doing at the time, and their state of mind; how anxious, somber and competent they felt, how focused they were on their fatigue and how demanding their current activity was. All three questionnaires contained the same six questions about how fatigued they felt at the time.

Items from the daily administered questionnaires were answered on a scale from 1 to 7. For example, the participant's quality of sleep was measured with the item 'I slept well' with one meaning I did not sleep well, and 7 meaning I did sleep well. All answers were filled in directly on the electronic diary.

Since the items themselves will be used as the variables of interest for this study, the usual test characteristics based on classical test theory are not particularly interesting. However, it may be necessary to consider the feasibility of the usage of electronic diaries with burn-out patients. This has been estimated as high (Sonnenschein, Sorbi, van Doornen & Maas, 2006). The compliance rate of the participants with burnout using the electronic diaries was high, with 81% of the alarmcontrolled diaries filled in, and respectively 94% and 96% for the evening diary and the morning diary.

Statistical analyses

The time series modeling was done with the MKFM6 program (Dolan, 2005)[1]. It can model stationary time series for one multiple subjects, and uses maximum likelihood estimation based on the Kalman Filter.

An exploratory vector auto-regressive model (VAR(1) model) was fitted for each subject (Figure 1). It includes three observed variables, respectively fatigue, energeticness and self-competence. One fixed regressor was included to control for the cyclic effects of the time of day; participants tend to be more tired in the morning and especially the evening, compared to the afternoon.

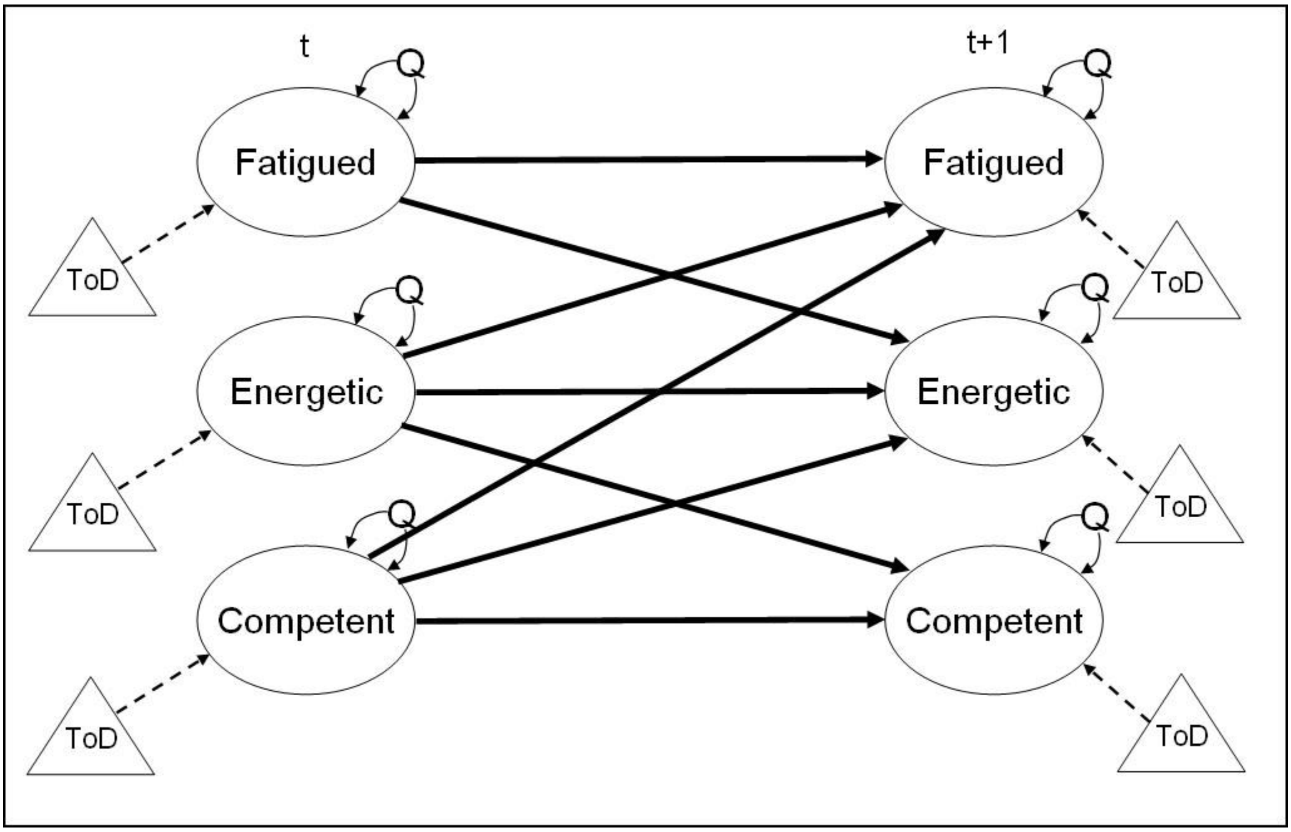


Figure 1. Fitted exploratory VAR(1) model. Round variables represent the variables of interest fatigue, energeticness and self-competence at time t and time t+1. Triangular variable represents fixed regressor time of day.

1 MKFM6 can be downloaded from http://users.fmg.uva.nl/cdolan/.

Results

In this study we focused only on the individuals with burn-out that took part in the first two weeks of measurement, and the other two weeks of measurement a half year later. This left 54 participants for the analyses. Of these participants 46.3 % was female and the mean age was 42.8 (sd 9.14). Most (62%) of the participants have had a upper-level education (dutch HBO or WO). The majority (96%) of the participants was currently employed, with an average of 33 to 40 working hours a week. Fifty percent of the participants has had complaints of fatigue for 22 months or longer.

The variables of interest for this study were those that reflected the symptomatology of burnout; respectively severe fatigue, and a cynical attitude towards work or work-related self-efficacy. Items that reflected fatigue were "I feel tired", "I feel exhausted", "I feel dead tired", and "I feel lethargic" [2]. Two items reflecting energeticness were "I feel energetic" and "I feel fit"[3]. Items that reflected self-efficacy or feelings of self-competence were "I feel competent right now", "What I'm doing right now I can handle well", "This activity is going well for me"[4].

These items were averaged in order to reflect three variables; respectively "fatigue", "energeticness" and "self-competence". This was done based on the content of the items and their intercorrelations (> .5 if added together). Next to data-reduction, aggregating the items had the advantage of normalizing the data. Many participants did not show the entire score range on the singular items, which is to be expected when an individual is measured for only two weeks. This does however have consequences for the continuity and normality of the participant's data, and continuously normally distributed data is one assumption of the time series analyses employed in MKFM6. Averaging the items resulted in more continuous and normal distributions for the variables of interest.

Before continuing with the time series models, some assumptions of these models should be considered. Assumptions of multivariate (stationary) time series that will be discussed are equal time intervals, multivariate normality and stationarity.

Equal time intervals means that between every time point (lag 1) the same amount of time should have passed. For instance, a participant could be measured every hour, or every week. If the time intervals are not equal, this can lead to under- or overestimation of the auto-covariances and cross-covariances, depending on the way the intervals differ. In this study two characteristics of the measurement process led to unequal time intervals. Firstly, one of the questionnaires was administered randomly throughout the day. The time intervals varied from about 1 to 5 hours. Secondly, there were periods of sleep after every day of measurement, which caused a large time

2 In dutch respectively: Ik voel me moe, uitgeput, bekaf, futloos.
3 In dutch respectively: Ik voel me energiek, fit.
4 In dutch: Ik voel me nu bekwaam, Wat ik nu doe kan ik goed aan, Deze activiteit gaat me goed af.

interval every night. The very large time intervals (>5 hours) were dealt with by adding missing observations to the data in those time intervals (1 for every three hours). This left a mean time interval of 2.3 hours (sd 1.1) over all subjects. Although this does not completely solve the issue of unequal time intervals, it does hopefully limit it.

Since the MKFM6 program uses maximum likelihood estimation, the dependent variables are assumed to be normally distributed for each participant. As mentioned before, items were averaged in order to get more continuous normal distributions for the variables of interest. However, a lot of participants still displayed very skewed and kurtosed distributions. Eight participants showed reasonably normal distributions, and another five had mostly problems of skewness in one of the three variables. Analyses were performed only for these thirteen participants. The rest of the participants showed severely skewed and kurtosed distributions for multiple variables and were deemed unfit for further analyses.

Stationarity means that the mean and variance of a time series should be constant over time. This means any trends or cycles in the time series should be controlled for. For instance, if a time series shows a decrease in mean depression score in the summer compared to winter, stationarity is violated. This trend should then be removed from the data; unless you are specifically interested in the trend itself, in which case you would specifically investigate the trend instead of controlling for it. In this study most subjects showed no obvious trends. Some evidence of a daily cycle however was visible in the time plots for some subjects, with the subject being more tired in the morning and evening than in the afternoon (See Figure 2 and 3 for the time series of each participant). This was modeled with a fixed regressor in the time series model (Figure 1).

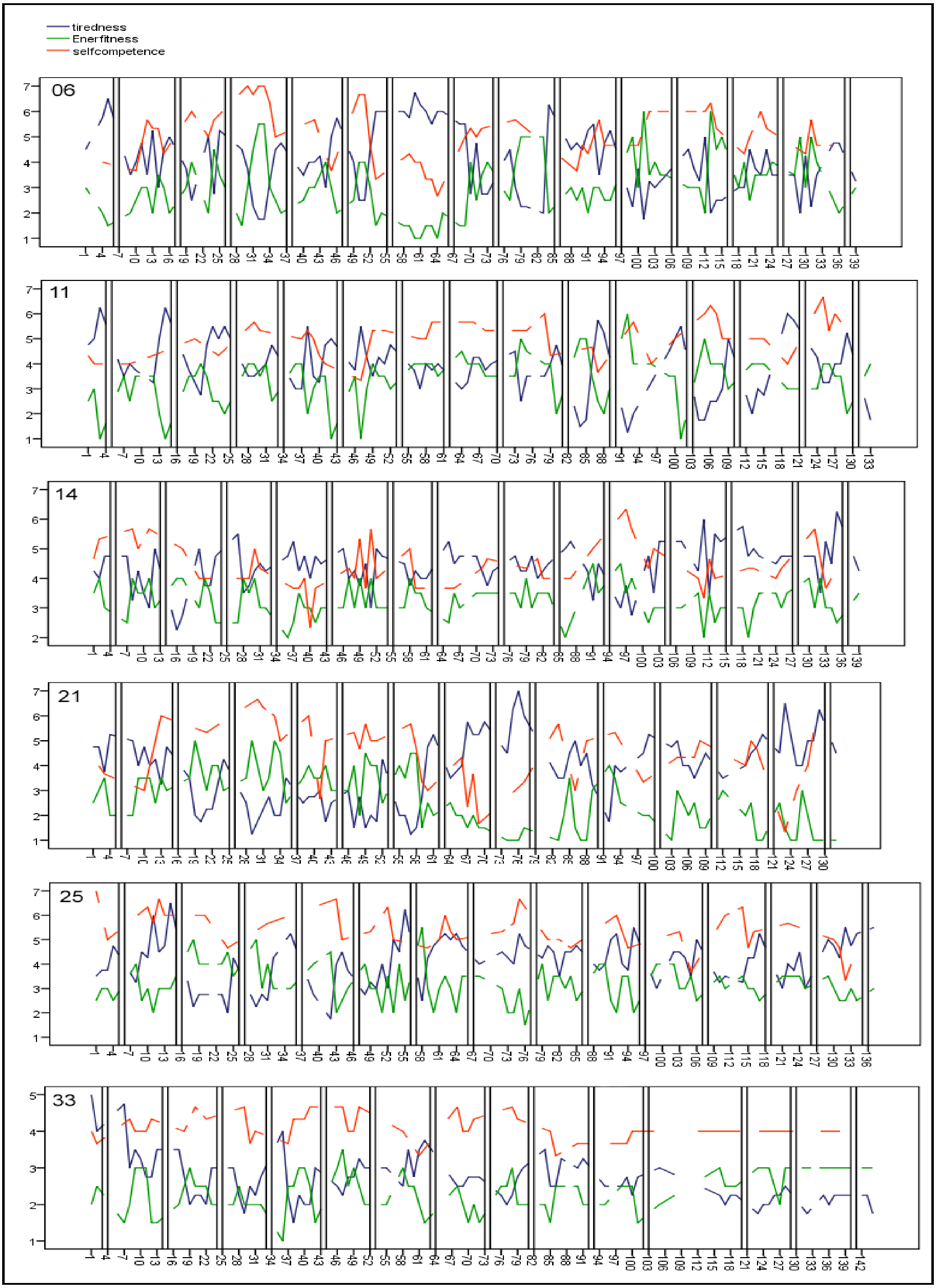


Figure 2. Time series for participant 06 to 33. Vertical lines represent the start of a new day. Diurnal patterns can be distinguished with convexes in energeticness and concaves in fatigue in the afternoon. It is also noticeable that the series for fatigue and energeticness are strongly negatively related at lag 0.

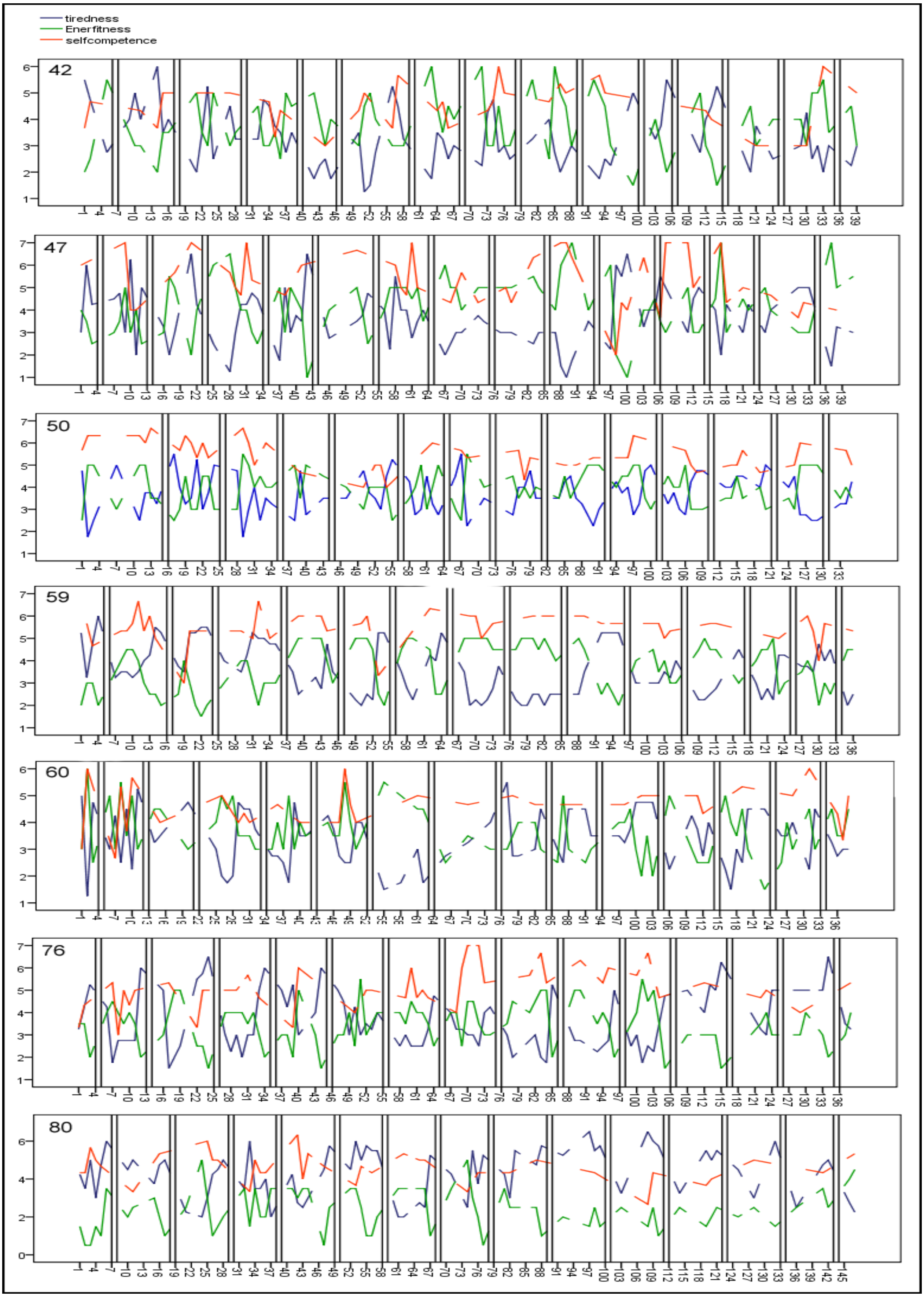
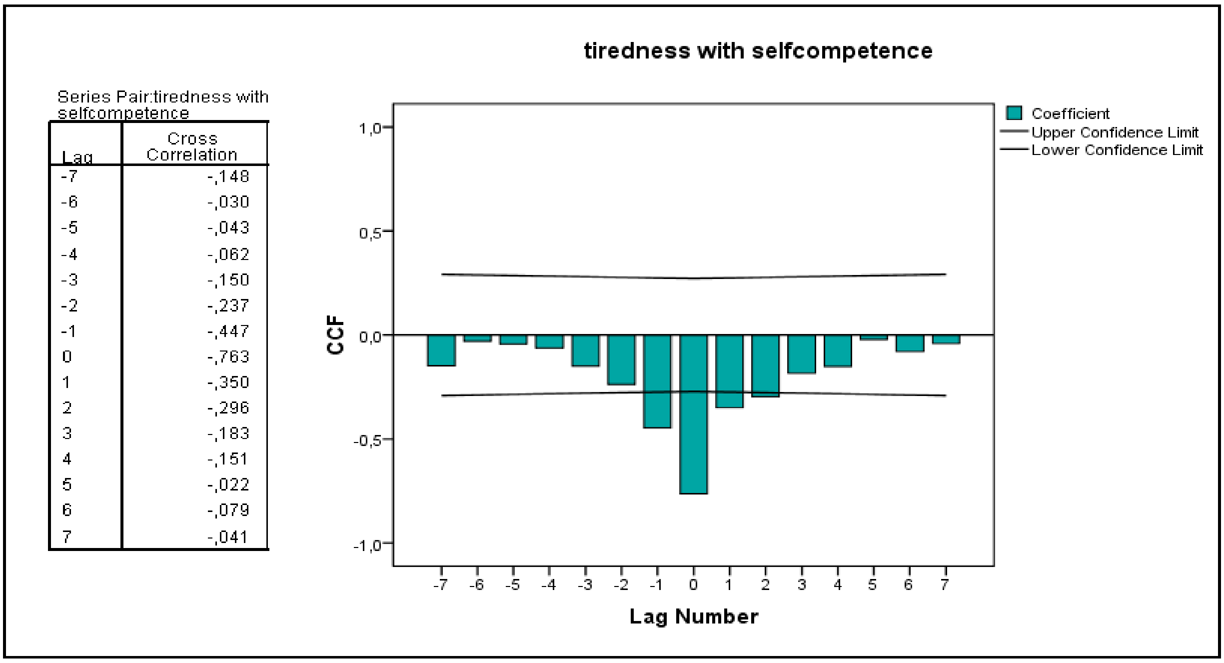


Figure 3. Time series for participant 42 to 80. Vertical lines represent the start of a new day. Diurnal patterns can be distinguished with convexes in energeticness and concaves in fatigue in the afternoon. It is also noticeable that the series for fatigue and energeticness are strongly negatively related at lag 0.

Q-graph auto-correlation and cross-correlation plots

Part of the internship was to write a new program to graphically represent the relationships between observed variables over time. I did this together with another intern (R. Hillen). The program was written in R and rests on the 'q-graph', a causal graph plotting R-package written by Sacha Epskamp[5].

We decided to describe the relationships between the observed variables over time with autocorrelations and cross-correlations. Auto-correlations and cross-correlations are calculated over time lags, and as such are directional. A lag is the distance between two time points The autocorrelations of a variable at lag 1 is then the correlation of the variable with itself from one time point to the next time point The autocorrelation at lag 2 would be the correlation of the variable with itself from one time point to the time point second in line after that. A cross-correlation is the correlation over a time lag between two variables. This means that for lag 1, there can be two cross-correlations: one for the first variable's correlation over time with the second variable, and one for the second variable's correlation over time with the first variable. In most cases these correlations are called the lag 1 and the lag -1 correlation for the two variables. Which correlation is the lag 1 and which is the lag -1 depends on the order in which you added the variables (See figure 4 for how cross-correlations are usually presented in programs, such as R or SPSS). Clearly, keeping track of all the auto- and cross-correlations and their directions over all lags of interest can be difficult when one wants to explore many different variables[6].


Series Pair:tiredness with selfcompetence

Lag	Cross Correlation
-7	-,148
-6	-,030
-5	-,043
-4	-,062
-3	-,150
-2	-,237
-1	-,447
0	-,763
1	-,350
2	-,296
3	-,183
4	-,151
5	-,022
6	-,079
7	-,041

Figure 4. Impression of how autocorrelations and cross-correlations are usually presented. These images were taken

5 Q-graph can be found online at http://sites.google.com/site/qgraphproject/about-qgraph .

6 The number of cross and autocorrelations equals (the number of variables ^{2} + the number of variables)* number of lags.

from spss 17.

In the program, cross-correlations and auto-correlations are presented in causal graphs. The nodes in the graph represent the observed variables, and arrows represent cross-correlations or autocorrelations. The size of the nodes indicates how central the observed variable is (how many connections it has), and the thickness of the arrows indicates the strength of the correlation. The thicker the arrow is, the stronger the correlation is. The color of the arrows indicates if it is a negative correlation (red) or a positive correlation (green). There remain a few issues that may be dealt with in the future. Firstly, the graphs do not take significance or standard errors of the correlations into account. Secondly, the program uses the crosscorrelation function of the standard stat package in R, which does not yet deal well with missings; it basically ignores missings altogether, which could result in overestimation or underestimation of correlations of series with lots of missings.

With the program it is possible to choose to present the correlations either for one lag (Figure 5), or for a number of lags (Figure 7). When the former option is chosen, a circular graph of the observed variables and their associations is presented. When the latter option is chosen, a chain-graph is presented, with between each two rows of observed variables the correlations for one lag.

A third option is to show the activations of the observed variables over time in a animated .gif (Figure 6). The activations are visible as color tones of the nodes. With each time-point the nodes change color to represent the variable's activation at that time point . The arrows in the graph are the correlations of one lag of choice (like in Figure 5).

Presenting the autocorrelations and cross-correlations in this manner provides a clearer overall picture of the relationships between variables over time than merely noting the cross-correlation, and at the same time can give an idea for possible models to be fitted exploratively.

In Figure 7 the crosscorrelations and autocorrelations between fatigue, energeticness and self-competence are presented for each participant for three lags. Interestingly, the participants each have a clearly different pattern of cross and autocorrelations (see appendix A). This illustrates the importance of individual analysis. A certain therapy may have been found to be successful for patients overall, the therapy will probably be more effective it is adjusted to suit personal needs. A model of the associations between key symptoms of the disorder can help to target focal points for treatment. For participant 47 for instance, it might be beneficial to focus on energizing her.

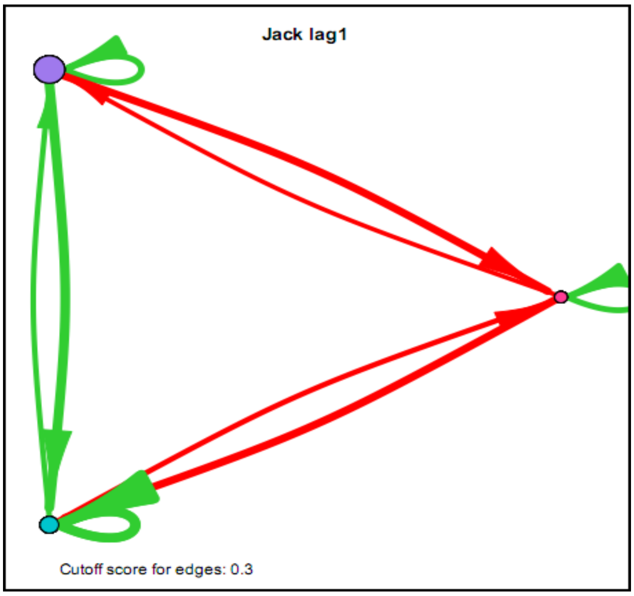


Figure 5. Example of the q-graph autocorrelation and cross-correlation plot. The nodes in the graph represent variable fatigue (pink), energeticness (purple) and self-competence (blue). The arrows represent autocorrelations (autoregressive arrows) and cross-correlations. A positive association is represented by a green arrow, a negative association is represented by a red arrow.

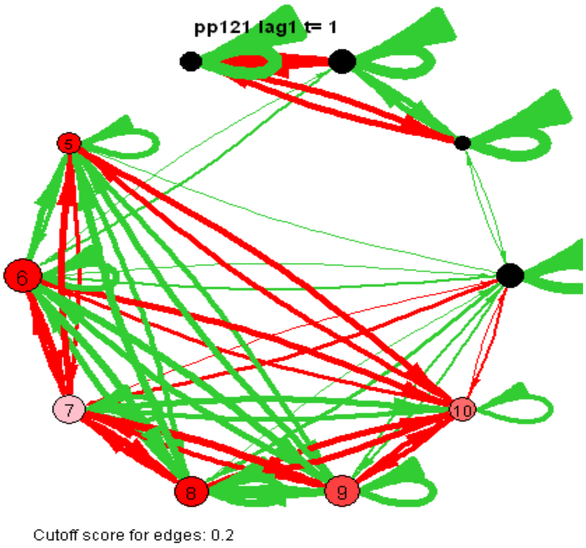


Figure 6. Example of a q-graph autocorrelation and cross-correlation animated .gif. The nodes in the graph represent items, their color tone represents the variables' activation. The darker red, the higher the activation. A black color tone indicates missingness at that time point The arrows represent autocorrelations (autoregressive arrows) and cross-correlations for one lag. A positive association is represented by a green arrow, a negative association is represented by a red arrow.

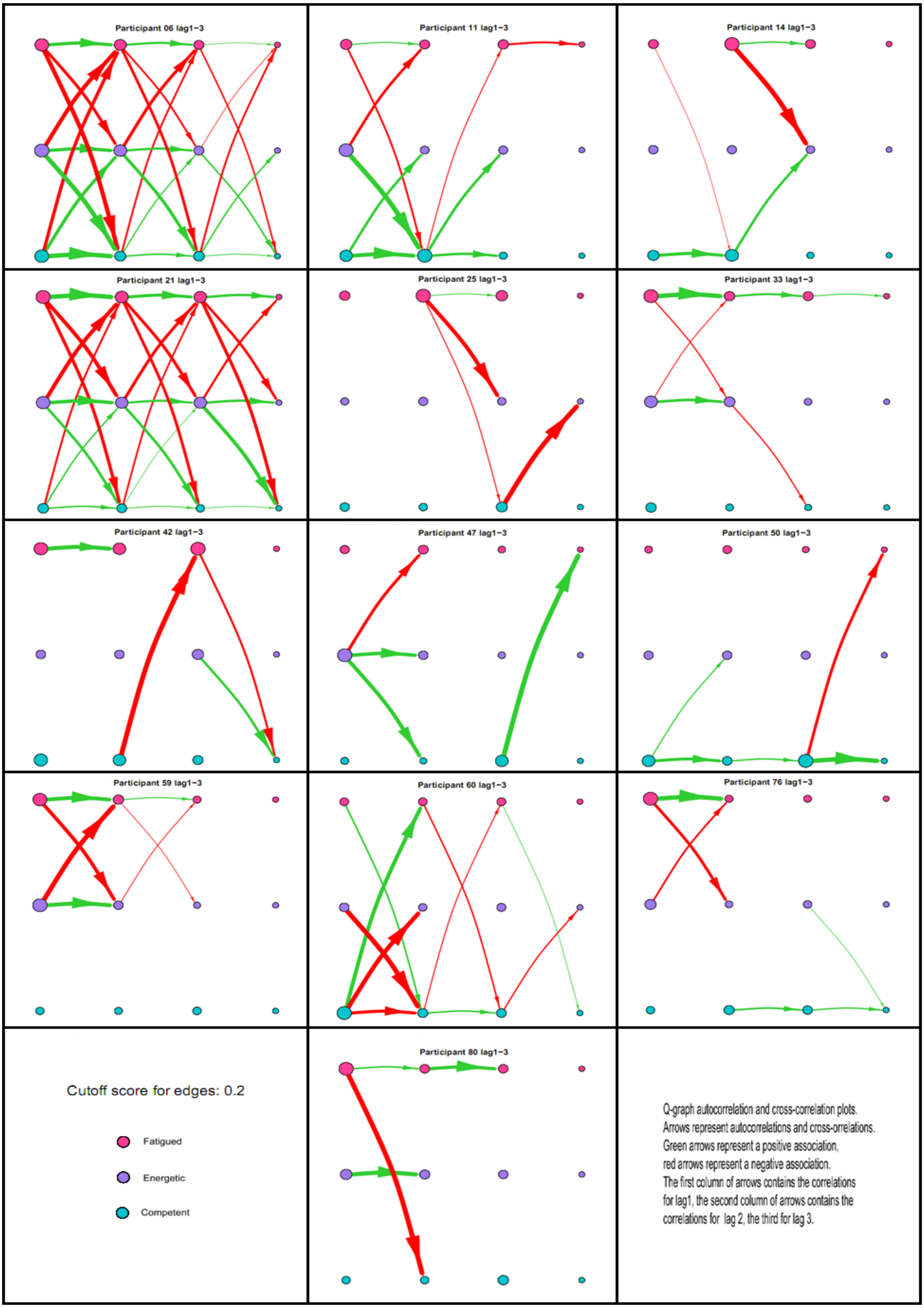
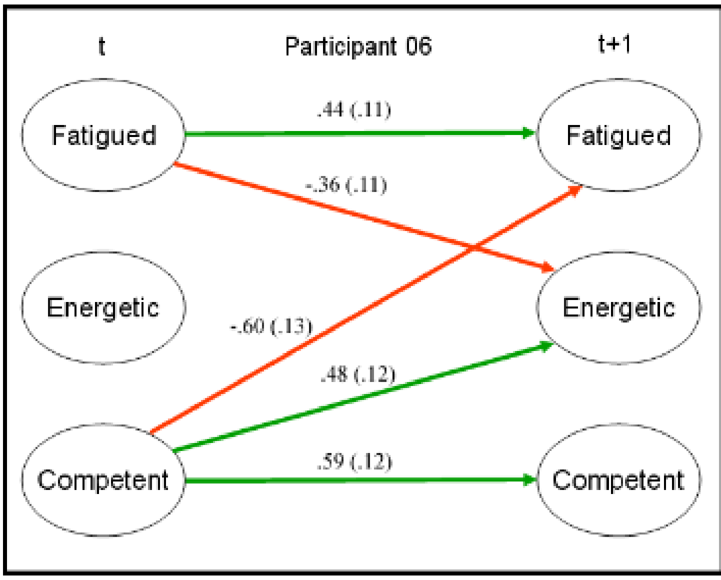


Figure 7. Q-graph autocorrelation and cross-correlation plots for all subjects.

Time series modeling for individuals

For participant 06 a model was found that seems sound. Participant 06 is male, age 40, and works in research. He had had complaints of fatigue since the last 6 years, the complaints had increased since the last four months. The time series length for participant 06 was 139, with 94 valid observations for fatigue (mean 4.04, sd 1.29) and energeticness (mean 3.09, 1.17), and 54 observations for self-competence (mean 5.1, sd 1.04).The model for this participant and the cross-correlation matrix can be found in Figure 5. The model seems interpretable. There is a negative effect of tiredness on energeticness one time point later. When the participant feels fatigued now, he is likely to feel un-energetic a few hours later. Self-competence has a negative effect on fatigue and a positive effect on energeticness. This means that if the participant feels self-competent now, he will feel less fatigued and more energetic later. Fatigue and self-competence are autorelated, which means that when the participant feels fatigued or competent now, he will probably also feel fatigued or competent later.

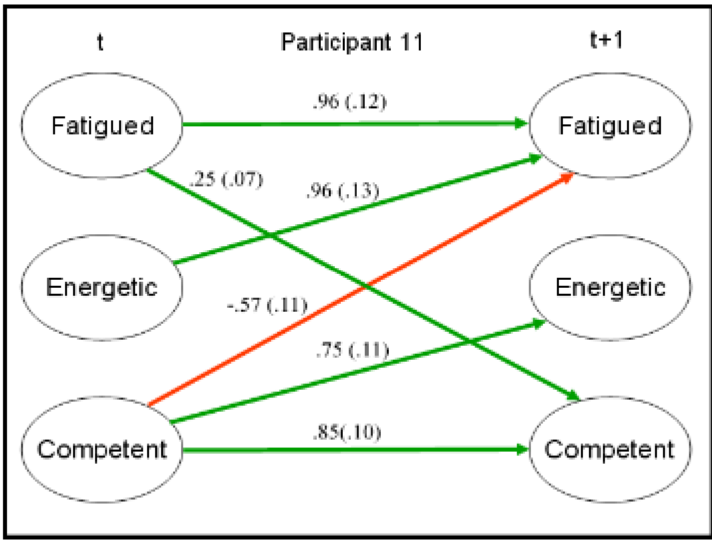


	fa.+1	en.+1	sc.+1
fa.	.38*(.09)	−.38*(.1)	−.35*(.14)
en.	−.36*(.1)	.37*(.09)	−.36*(.14)
sc.	−.45*(.14)	.44*(.14)	.48*(.10)

Figure 8. Time series model and cross-correlation matrix for participant 6.

The time series modeling harbored some unexpected results (See Figure 12 and appendix A for results for all participants). First it should be mentioned that most participants showed little or no significant cross-correlations and autocorrelations. This provides little basis for fitting models, and could be the cause of strange results for these participants. However, for participants that did show some significant auto- and cross-correlations, models were found that were not as expected as well. This will be illustrated with two examples.

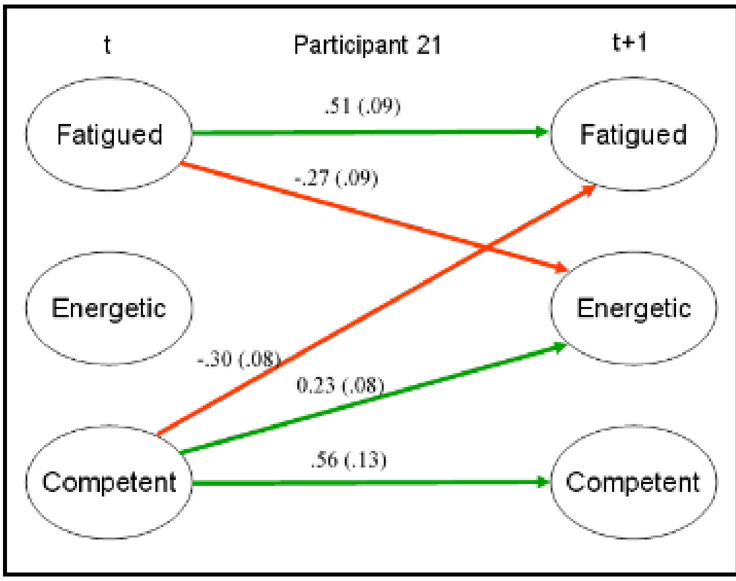
Participant 11 had a series length of 133 observations (including missings). For energeticness (mean 3.89, sd 1.12) and fatigue (mean 4.41, sd 0.98) 81 observations were valid, for self-competence (mean 5.03, sd 0.76) 43 observations were valid. The found model and cross-correlation matrix for participant 11 can be found in Figure 3. There are a few things that catch the attention when inspecting participant 11's model. Firstly, a positive effect of energeticness on fatigue is found. This means that the more energetic the participants is now, the more tired the participant will be a moment later. This seems counter-intuitive since fatigue and energeticness are expected to be opposites each other. A look at the cross-correlation matrix confirms the expectation; energeticness is negatively crosscorrelated with fatigue. Secondly, the model shows a negative effect of self-competence on fatigue, but a positive effect of fatigue on self-competence. This would mean that the more self-competent the participant feels, the more tired the participant will feel a moment later, and the more tired the participants feels, the less competent the participant will feel another moment later. Again this seems counter-intuitive, and the cross-correlations show no significant effect from fatigue on self-competence. Finally, in the model self-competence influences both fatigue and energeticness while there is no evidence for these relationships judging from the cross-correlations. Using the crosscorrelations however may not be a foolproof way to check for the plausibility of the models.



	fa.+1	en.+1	sc.+1
fa.	.39*(.10)	−.37*(.11)	−.21(.15)
en.	−.24*(.11)	.25*(.10)	.29(.15)
sc.	−.21(.21)	.18(.15)	.28*(.10)

Figure 9. Time series model and cross-correlation matrix for participant 11.

Participant 21 had time series of length 132, with 90 valid observations for fatigue (mean 3.76, sd 1.36) and energeticness (mean 2.67, sd 1.15) and 44 valid observations for self-competence (mean 4.41, sd 123). The model for this participant and the cross-correlation matrix can be found in Figure 4. At a fist glance, the model does not seem insensible. Fatigue and self-competence are positively related to themselves in the future; the more fatigued or self-competent the participant feels, the more she will in the future. A negative relationship between fatigue and energeticness seems to make sense as well. Self-competence negatively affects fatigue, while it positively affects energeticness. In other words, the more self-competent the participant feels, the more energetic and less fatigued she feels. Self-competence seems the most influential variable in the model. Yet, the cross-correlations show no significant correlations for self-competence. Judging from the cross-correlations, self-competence should have no effect on the other variables, and energeticness should affect fatigue.



	fa.+1	en.+1	sc.+1
fa.	.65*(.09)	−.51*(.11)	−.15(.15)
en.	−.56*(.11)	.52*(.09)	.13(.15)
sc.	−.23(.15)	.11(.15)	.19(.10)

Figure 10. Time series model and cross-correlation matrix for participant 21.

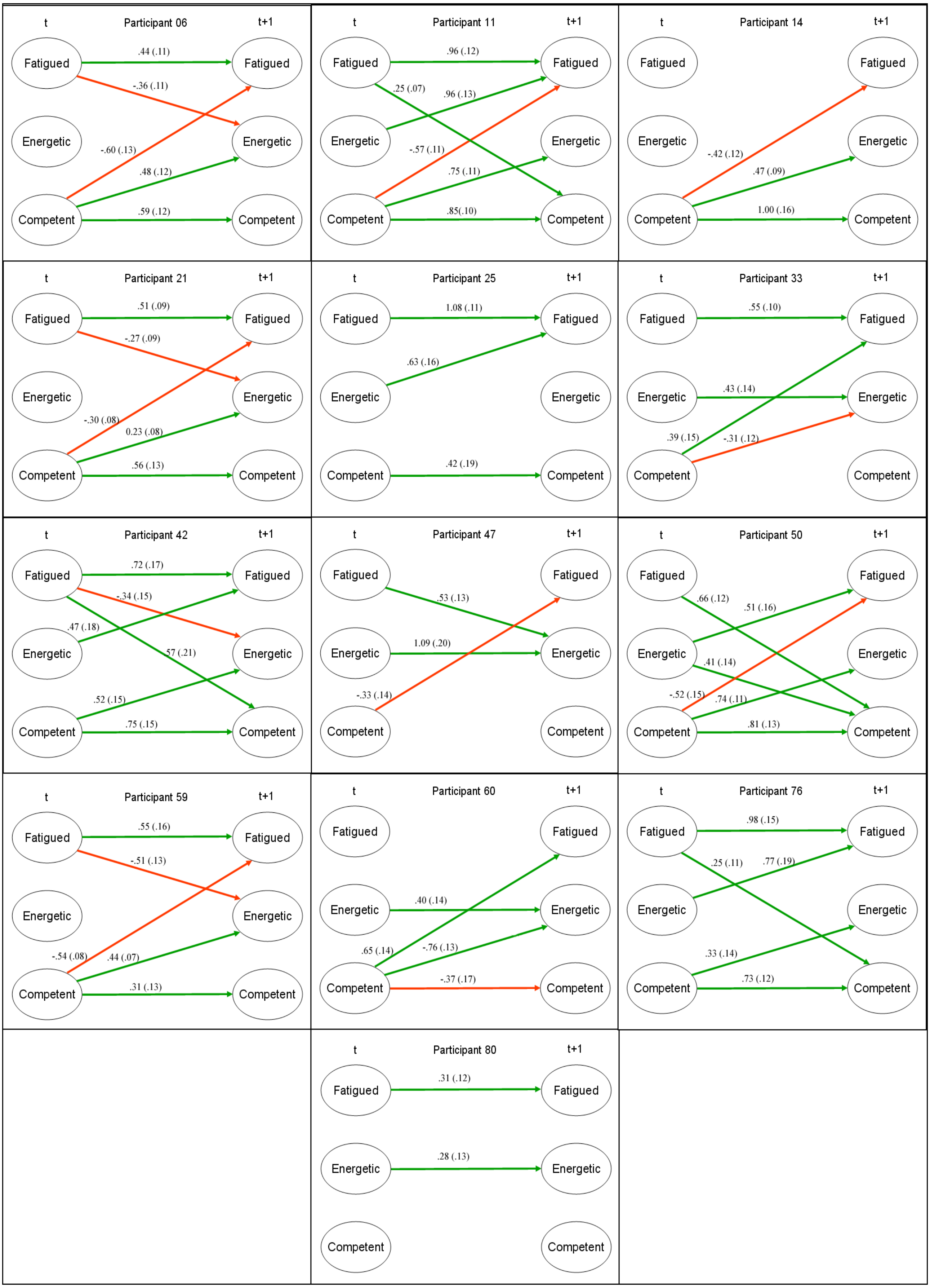


Figure 12. Time series model for every participant.

Time series model for all participants

Next to explorative time series analyses for each individual participant, a time series model was fitted for all participants as a group. The same basic model as before was employed (see figure 1), with equal parameters estimated for all the participants, except for the variable means which were allowed to differ over participants. The model found shows a positive auto-relationship over time for fatigue and for self-competence. Self-competence had a positive effect over time on energeticness, the more self-competent the participants feel (overall) at one time point, the more the participants will feel energetic at the next time point. Fatigue has a positive effect over time one self-competence; the more fatigued the participants (over all) feel at one time point the more competent the participants feel at the next time point. This last result does not seem very sensible, it would be expected that fatigue would have a negative influence on how competent someone feels. However, strange results in this analysis should not come as a surprise. Although this group analysis should have more power than the individual analyses, the issues that may have played a part in the individual analyses may play a part in the group analysis. Furthermore, for the group analysis it is assumed that all participants have (approximately) the same underlying model. If the participants in fact have very different models, which seems to be the case, the results for the group might not be representative for any individual in that group, and perhaps might not be sensible. In the future however the variability of the individuals can be taken into account by using random effects.

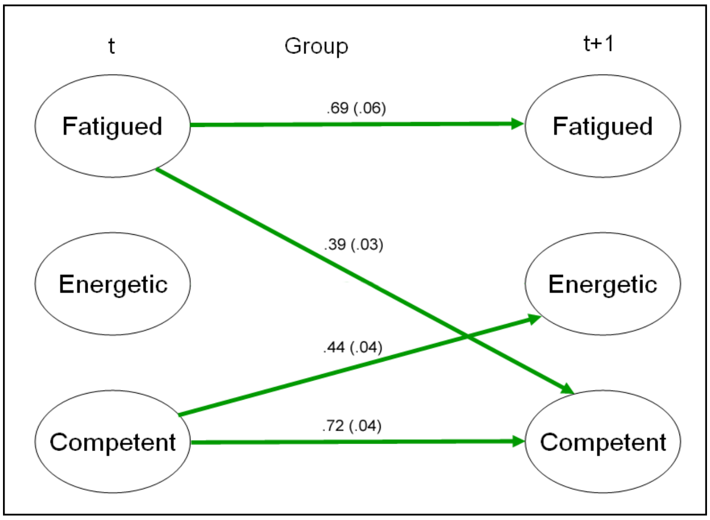


Figure 12. Time series model for all the participants as a group.

Discussion

This is the first time that psychopathological data is studied in this way, which has not been trivial. The process has taught us a lot. Taking the time to explain the unexpected results could benefit future studies. In explaining the unexpected results, many different circumstances should be considered. The data used for this study was not collected with time series modeling in mind, which is not without consequences. It was mentioned before that most participant showed little significant auto- and cross-correlations. Cross and auto-covariances or correlations are the basis for time series modeling, so if there is little cross- and auto-covariance present it can make it difficult to fit sensible models.

Small and non-significant auto and cross-correlations can be the results of many things. The time series collected were short for the variables fatigue and energeticness (around 90 valid observations per participant). The self-competence variable was only measured with the random questionnaire during the day and included missings, leaving even shorter time series for this variable (around 45 valid observations per participant). Next to that there were little observations, missings were added to the time series in order to ensure more stationary series; however this causes more 'holes' in the data which could also leave less data for the auto- and cross-correlations. These issues have probably led to little power, which makes it difficult to find significant associations if these associations exist in the population.

It should also be considered that the time-intervals chosen may have been too large to pick up strong associations. If causal systems occur on a small timescale, they might not be picked up anymore when measurement takes place on a larger timescale. For instance, if the causes and effects take place within a half hour of each other, the effect may not be picked up if you measure once every day. The time intervals for this study was about 3 hours, which is already a small time interval considering practical circumstances.

A problem that coincides with the problem of low power, is that there was just little variation within the variables to begin with. The root of this problem could be either in the data gathering or in the actual underlying constructs and their actual relationships, or both.

The time-intervals for the time-series were unequal. Relatively large time-intervals may have deflated the auto- and cross-covariances. Furthermore, most participants did not show the full score-range on the individual items. With analyses relying on multiple subjects such as repeated measures ANOVA, this is not necessarily a big problem because different subjects will show different scores, ensuring a wider score-range over participants. With analyses for one participant however, this is certainly something worth consideration. Two weeks of following the participant may not be nearly long enough for the participant to show all of the score-range, especially when the variables of interest are expected to be strongly autorelated, resulting in little variance.

Furthermore, the participants usually had burnout complaints for many months or even years. It could be that these participants symptoms show little variability because their causal system has become stable. That the invariance of these symptoms may be characteristic of chronic burnout is also expressed by Sonnenschein et al. (2007), who hypothesize that "[...] the end-stage, clinical burnout may be characterized by a continuous state of exhaustion showing little variability and flattened diurnal patterns.". They found that the within variability of feeling tired (but not of feeling exhausted) was smaller for individuals with burnout than the within variability of healthy controls. They also found flattened diurnal patterns of tiredness for individuals with burnout compared to healthy controls. So there very well may be a stable causal system at work here, yet this system cannot be found with statistical modeling due to the invariance that is characteristic of the system itself.

For some participants multiple significant autocorrelations and cross-correlations were found. For one participant this seemed to deliver interpretable results, but for others it did not. It seems unlikely that the problems with the models have arisen purely due to a lack of power and variability. It is likely that the strange models have arisen due to some kind of misfit that disturbed the models. Where the misfit comes from exactly is not easy to track.

It is possible that some misfit has arisen from violations of the assumptions of time series models. Some participants included in the analyses showed skewed distributions for a variable. This could have influenced the maximum likelihood estimations of the parameters, especially with small samples. In SEM the estimates seem to get overestimated when distributions are medium to highly skewed (Cuttance & Ecob, 1987). This may explain some of the unexpected significant parameters in some of the models (see participant 21, Figure 4). However, it is unclear if these effects also hold for time series.

It is also possible that misfit has arisen due to associations between the variables that have not been modeled. It may be that misfit resulted from only modeling lag 1 associations and not more. Furthermore, multicollinearity (at lag 0) of the variables energeticness and fatigue may have played a part for some participants. Misfit caused by the associations between these two variables may have spilled over to the rest of the model. For example, for participant 11 (Figure 3) energeticness and fatigue were highly multicollinear at lag 0 with $r = -.83$ (sd .11). One of the counterintuitive results for participant 11 was that a positive effect from energeticness on fatigue. The correlations a negative relationship between energeticness and fatigue. Another surprising results was the positive relationship between fatigue and self-competence, while no significant correlations are found. When inspecting the results we also find a significant positive correlation from energeticness on self-competence, but not a significant relationship in the model. Perhaps the missing relationship between energeticness and self-competence is made up for through a positive relationship between energeticness and fatigue, and in turn a positive relationship between fatigue and self-competence. In other words, the associations of energeticness may be modeled through fatigue, perhaps because of their strong multicollinearity. Similar effect were found when fixed regressors for aspects of sleep were added to the model exploratively (Appendix B).

Conclusions

Fitting time series models for individuals with burnout proved to not be a trivial task; for one participant an interpretable model was found, the models of other participants seemed less interpretable. For the one participant, self-competence seems to play the most important role in his causal model for burnout. If his self-competence increases, it is likely that he will become less fatigued and more energetic. These kind of results could be valuable for therapists; in this case perhaps the focus could lie on getting the participant to feel more self-competent.

The limitations of this study can provide recommendations for the future. First and foremost, long time series should be collected, preferably with no missing data. Although time series can handle missing data relatively well, it is still detrimental for the power of the analyses. In order to prevent little variation in the measured variables of interest, it may be important to collect the time series over a long period of time as well, instead of just collecting long series.

Time intervals should preferably be of equal size between each time point. They should be spaced in a way so that the associations between variables can be captured, but the repeated measurements should also interfere the least as possible in a participants life. For example, if one is primarily interested in the influence of aspects of sleep on the participant's feelings of anxiety, it may be only necessary to take one daily measure. If one is primarily interested in the effects of moments of stress on feelings of depression, it may be necessary to measure the participant more often.

Participants should be instructed carefully on how to fill in questionnaires. Answers should be given relative to how the participant have answered before, in the period of measurement; not in relation to other persons, or in relation to when they were healthy. This could perhaps ensure more spread in the participant's answers.

The next thing to consider are which participants to measure. For example, it is important to consider if participants should be measured that are at risk for a disorder, that are at the early onset of a disorder, or participants that have a full fledged disorder. Which phase is chosen will likely relate to, for instance, the variability of the variables. Maybe the participant should be followed throughout all these phases, in which case a trend could be modeled for the participant.

Ultimately, the knowledge gained from time series studies of psychopathology could be greatly beneficial for both science and practice, supplying new knowledge about psychological disorders and new directions for psychotherapy. Future studies of psychopathology employing time series could be much facilitated by taking the issues mentioned in this study into account.

References

Arrindell W.A., & Ettema J.H.M., (2002). Symptom Checklist SCL-90: handleiding bij een multidimensionele psychopathologie-indicator [Symptom Checklist SCL-90: manual on a multidimensional psychopathology indicator]. The Netherlands: Swets Test Publishers.

Blonk, R., Brenninkmeijer, V., Lagerveld, S., & Houtman, I. (2006). Return to work: A comparison of two cognitive behavioral interventions in cases of work-related psychological complaints among the self-employed. *Work & Stress, 20*, 129-144.

Borritz, M., Rugulies, R., Christensen, K. B., Villadsen, E., & Kristensen, T. S. (2006). Burnout as a predictor of self-reported sickness absence among human service workers: Prospective findings from three year follow up of the PUMA study. *Occupational and Environmental Medicine, 63*, 98-106.

Borsboom, D. (2008). Psychometric perspectives on diagnostic systems. *Journal of Clinical Psychology, 64*, 1089-1108.

Bültmann U., De Vries M., Beurskens A. J. H. M., Bleijenberg G., Vercoulen J. H., & Kant I, (2000). Measurement of prolonged fatigue in the working population: determination of a cutoff point for the Checklist Individual Strength. J Occup Health Psychol, 5, 411-416.

Chatfield, C. (2004). The Analysis of Time Series: An introduction. Boca Raton: Chapman & Hall, CRC Press.

Cuttance, P., & Ecob, R. (1987). Structural modeling by example: applications in educational, sociological, and behavioral research. Cambridge: Cambridge University Press.

Dolan, C. V. (2005). MKFM6: Multi-group, multi-subject stationary time series modeling based on the Kalman fillter.

Houtman, I. L. D., Schaufeli, W. B., & Taris, T. (2000). Psychische vermoeidheid en werk: cijfers, trends en analyses [Psychological fatigue and work: figures, trends and analyses]. Alphen aan de Rijn: NWO/Samsom.

Kant, I., Jansen, N. W. H., Van Amelsvoort, L. G. P. M. , Mohren, D. C. L., & Swaen, G. M. H. (2004). Burnout in de werkende bevolking. Resultaten van de Maastrichtse Cohort Studie [Burnout in the working population. Results of the Maastricht Cohort Study]. *Gedrag & Organisatie, 17*, 5-17.

Kim, N. S., & Ahn, W.(2002). Clinical psychologists' theory-based representations of mental disorders predict their diagnostic reasoning and memory. *Journal of Experimental Psychology, 131*, 451-476.

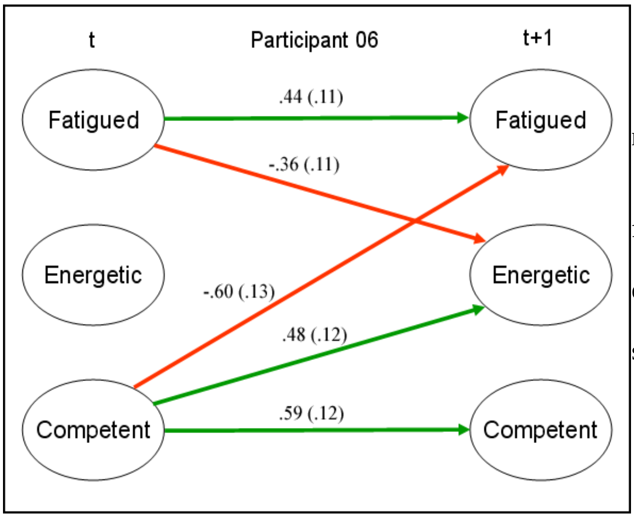
Maas, H. L. J., van der, Dolan, C., Grasman, R. P. P. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review, 113,* 842-861. Schaufeli WB,

Van Dierendonck D. Utrechtse Burnout Schaal: Handleiding [Utrecht Burnout Scale: Manual]. Lisse, The Netherlands: Swets Test Publishers, 2000.

Sonnenschein, M. A., Sorbi, M. J., van Doornen, L. J. P., & Maas, C. J. M. (2006). Feasibility of an electronic diary in clinical burnout. *International Journal of Behavioral Medicine , 13,* 315-319.

Sonnenschein, M., Sorbi, M. J., van Doornen, L. J. P., Schaufeli, W. B., & Maas, C. J. M. (2007). Electronic diary evidence on energy erosion in clinical burnout. J*ournal of Occupational Health Psychology, 12,* 402-413.

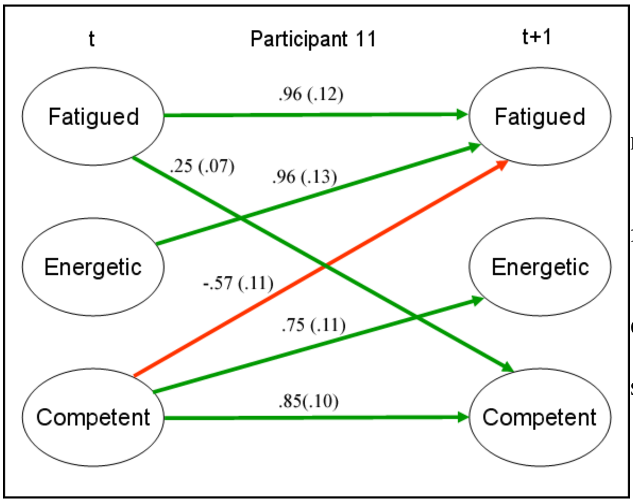
Appendix A. Results for individual participants



Logl: -198.62
-2Logl: 397.24

	fa.	en.	sc
mean	4.04 (1.28)	3.09 (1.16)	5.11 (1.03)

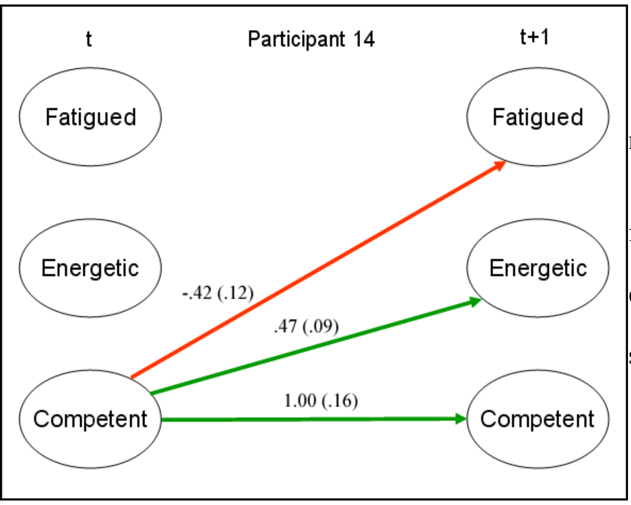
	fa.+1	en.+1	sc.+1
fa.	.38*(.09)	–.38*(.1)	–.35*(.14)
en.	–.36*(.1)	.37*(.09)	–.36*(.14)
sc.	–.45*(.14)	.44*(.14)	.48*(.10)



Logl: -122.371
-2Logl: 244.743

	fa.	en.	sc
mean	3.89 (1.15)	3.41(.97)	5.03 (.75)

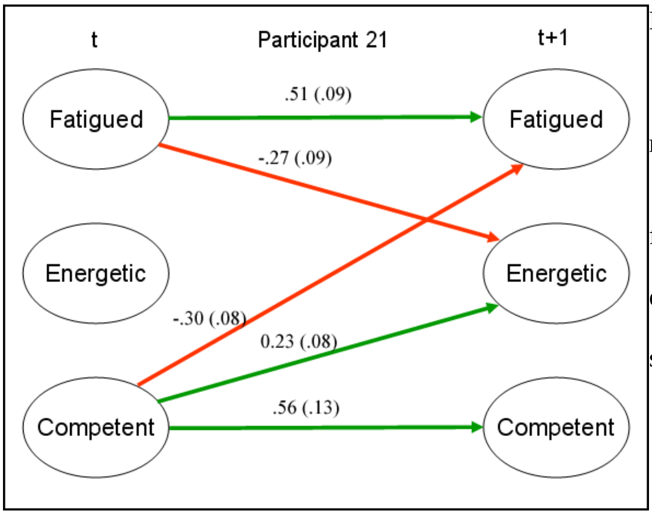
	fa.+1	en.+1	sc.+1
fa.	.39*(.10)	–.37*(.11)	–.21(.15)
en.	–.24*(.11)	.25*(.10)	.29(.15)
sc.	–.21(.21)	.18(.15) .28*(.10)	



Logl: -96.84
-2Logl: 193.69

	fa.	en.	sc
mean	4.38 (.74)	3.27 (.57)	4.49 (.76)

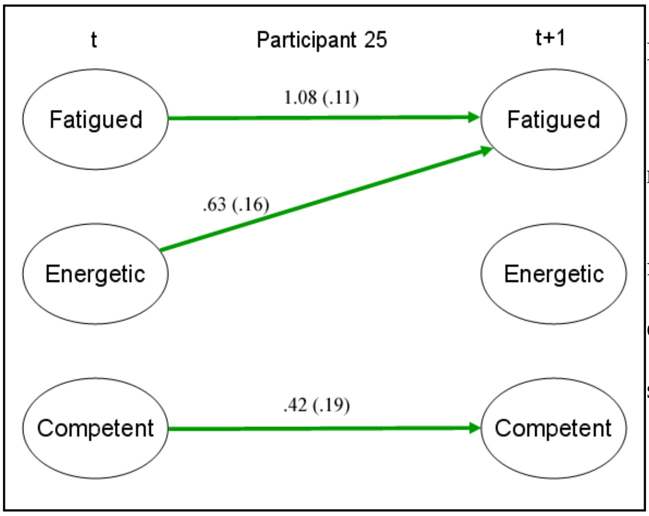
	fa.+1	en.+1	sc.+1
fa.	.05(.09)	.06(.10)	.01(.15)
en.	–.05(.11)	.06(.09)	–.07(.15)
sc.	–.09(.15)	.16(.15)	.15(.11)



Logl: -172.68

-2Logl: 345.37

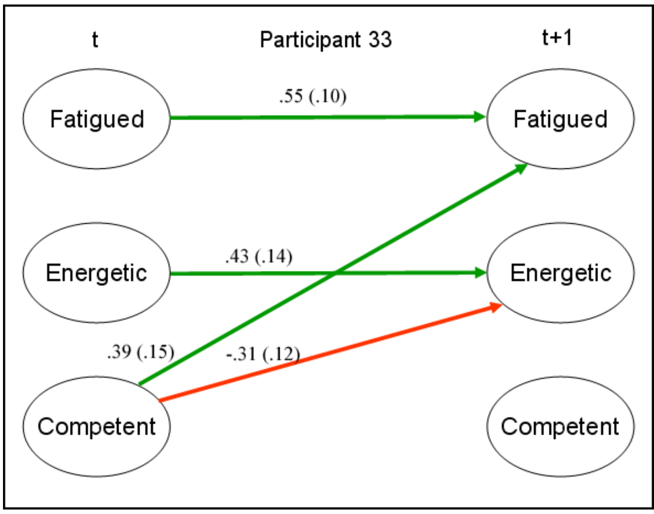
	fa.	en.	sc
mean	3.76 (1.35)	2.67 (1.16)	4.41 (1.22)
	fa.+1	en.+1	sc.+1
fa.	.65*(.09)	–.51*(.11)	–.15(.15)
en.	–.56*(.11)	.52*(.09)	.13(.15)
sc.	–.23(.15)	.11(.15)	.19(.10)



Logl: -142.76

-2Logl: 285.39

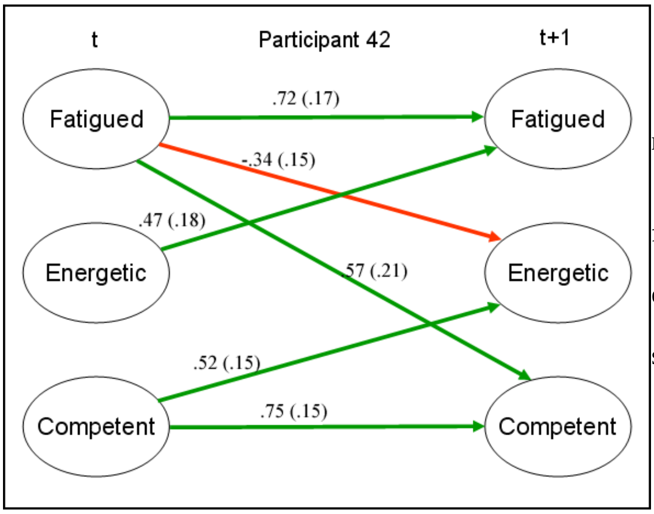
	fa.	en.	sc
mean	4.13 (.98)	3.16 (.76)	5.38 (.81)
	fa.+1	en.+1	sc.+1
fa.	.29*(.09)	–.06(.11)	.06(.16)
en.	–.15(.11)	.04(.09)	–.08(.16)
sc.	.07(.16)	–.10(.16)	.18(.1)



Logl: -5.44

-2Logl: 10.87

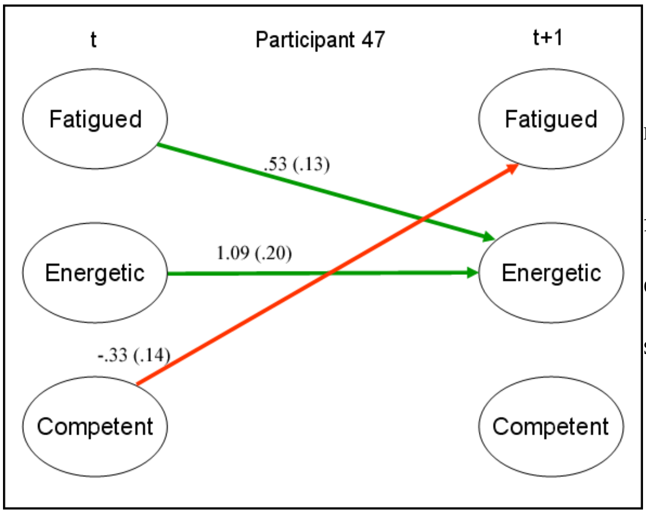
	fa.	en.	sc
mean	2.65 (.64)	2.40(.53)	4.09(.34)
	fa.+1	en.+1	sc.+1
fa.	.29*(.09)	.06(.11)	.06(.16)
en.	–.15(.11)	.04(.09)	–.08(.16)
sc.	.07(.16)	–.10(.16)	.18(.1)



Logl: -152.43
-2Logl: 304.86

	fa.	en.	sc
mean	3.31 (1.10)	3.78 (1.09)	4.52 (.80)

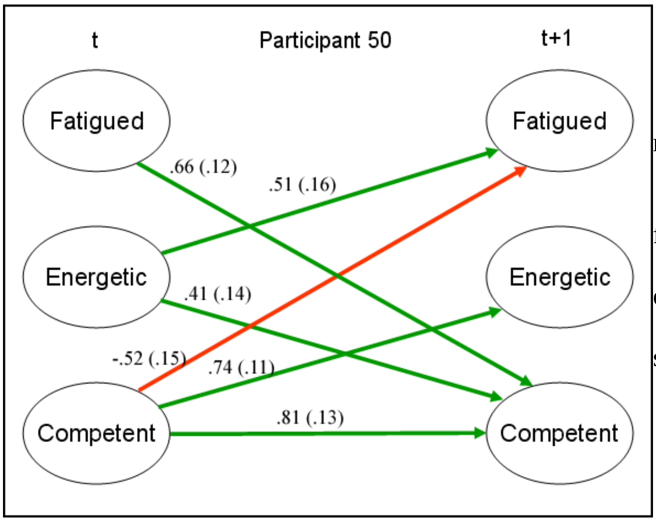
	fa.+1	en.+1	sc.+1
fa.	.30*(.1)	–.15(.12)	.06(.18)
en.	–.18(.18)	.14(.09)	.03(.18)
sc.	–.16(.18)	.17(.18)	.01(.11)



Logl: -224.29
-2Logl: 448.58

	fa.	en.	sc
mean	3.69 (1.28)	4.23 (1.3)	5.32(1.2)

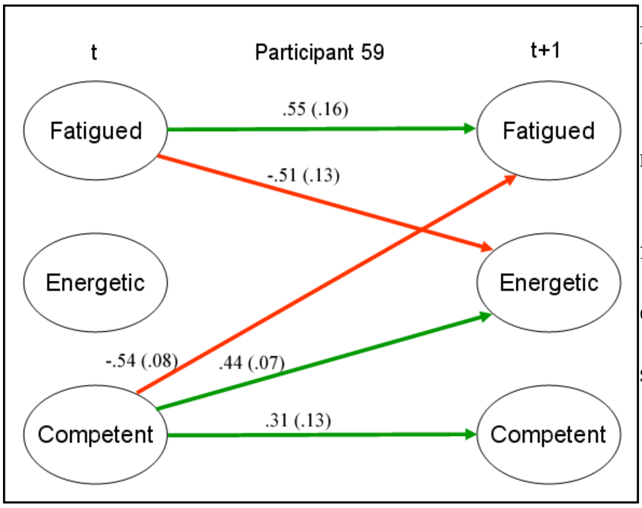
	fa.+1	en.+1	sc.+1
fa.	.46*(.09)	–.26*(.11)	–.09(.14)
en.	.37*(.11)	.33*(.1)	.12(.14)
sc.	–.11(.14)	.12(.14)	.11(.1)



Logl: -119.14
-2Logl: 238.27

	fa.	en.	sc
mean	3.69 (.89)	3.99 (.79)	5.44(.65)

	fa.+1	en.+1	sc.+1
fa.	.08(.09)	–.17(.11)	.08(.15)
en.	–.17 (.11)	.09(.09)	.04(.15)
sc.	–.09(.15)	.06(.15)	.16(.11)

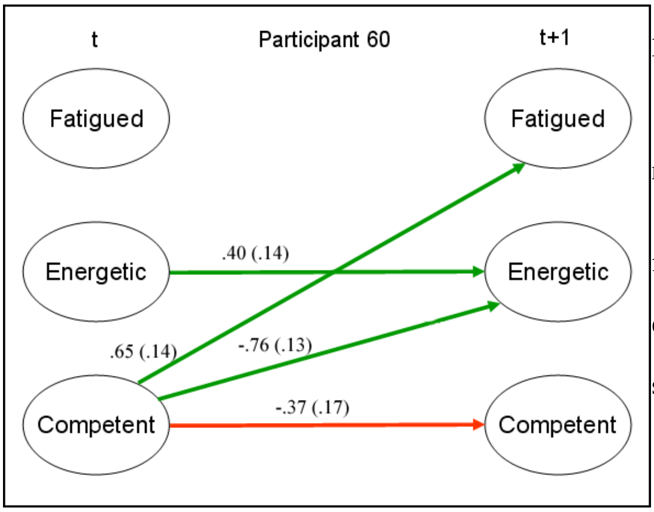


Logl: -124.8

-2Logl: 249.61

	fa.	en.	sc
mean	3.6 (1.14)	3.74 (1.05)	5.45 (.69)

	fa.+1	en.+1	sc.+1
fa.	.59*(.09)	–.60*(.11)	.02(.15)
en.	–.57*(.11)	.58(.09)	.06(.15)
sc.	–.14(.15)	.12(.15)	–.03(.1)

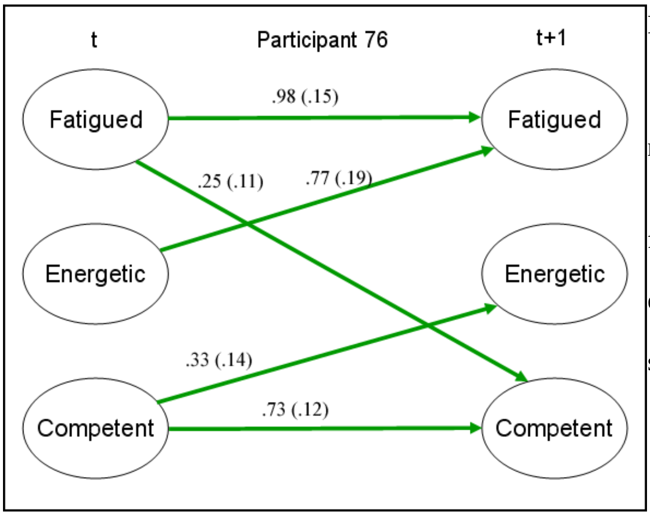


Logl: -132.12

-2Logl: 264.23

	fa.	en.	sc
mean	3.39 (1.05)	3.69 (.89)	4.6 (.59)

	fa.+1	en.+1	sc.+1
fa.	.03(.09)	–.02(.11)	.22(.17)
en.	.12(.11)	–.07(.09)	–.20(.17)
sc.	.22(.17)	–.38*(.17)	–.28*(.11)

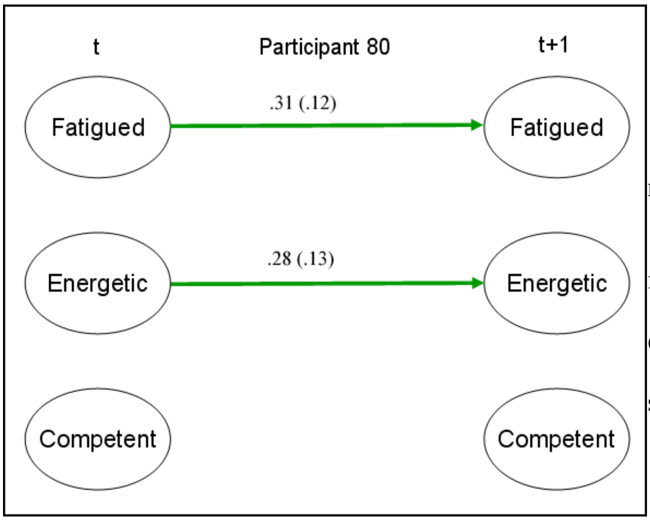


Logl: -183.70

-2Logl: 367.41

	fa.	en.	sc
mean	3.65 (1.27)	3.52 (.95)	5.09 (.94)

	fa.+1	en.+1	sc.+1
fa.	.19*(.09)	.28*(.11)	.02(.14)
en.	–.18(.11)	.33*(.09)	.14(.10)
sc.	.06(.14)	.11(.14)	–.01(.1)



Logl: -142.33
-2Logl: 184.66

	fa.	en.	sc
mean	4.31 (1.28)	2.39 (1.07)	4.57 (.77)

	fa.+1	en.+1	sc.+1
fa.	.26*(.09)	–.11(.12)	–.06(.17)
en.	–.19(.12)	.24*(.09)	–.06(.17)
sc.	–.24(.17)	–.09(.17)	.01(.11)

Appendix B Fixed regressors

At the start of the analyses it was tried to include fixed regressors for different characteristics of sleep in the model. Participants filled in a questionnaire after waking, including questions about how they slept the last night. Items that covered this were “I slept well”, “I had trouble with falling asleep”, “I woke … times last night” and “I slept for … hours”.[7] Since the participants only sleep once a day, only 12 to 14 observations were available for the sleep characteristics. It was decided to impute the missing values per day with the sleep observations of the morning before. So, each observation on one day will have the same value. Since it is likely that how a person slept the night before influences the total following day, this seems defensible. This does however severely inflated the autocorrelations for these variables, which can also result in inflated cross-correlations (Chatfield, 2004). Because of this problem, it was chosen to add the variables as fixed regressors rather than as a variable modeled over time.

Including the fixed regressors however seemed to cause misfit in the time series models. The fixed regressors had counterintuitive relations with the variables modeled over time. Te variables modeled over time also had counterintuitive relations with each other This was later found to be mostly due to multicollinearity of the fixed regressors. With using only one fixed regressor (“I slept well) the counterintuitive results decreased. However, the one fixed regressor seemed to also cause misfit by itself. When in the model associations of the fixed regressor with the time-modeled variables were not found that were indicated by (lag 0) correlations, this seemed to lead to strange associations between the time-modeled variables. In the light of all these results, it was decided to refrain from modeling the sleep variables. The counterintuitive associations in the models were decreased, but were not completely diminished.

7 In Dutch respectively: “Ik heb goed geslapen”, “Ik had moeite met inslapen”, “Ik ben vannacht … keer wakker geworden” and “Ik heb vannacht … uur geslapen”.

Mkfm6 How-to v0.0

This document is made as an addition to the mkfm6 manual. It provides some practical instructions and tips for mkfm6 that may not be obvious from the manual. This includes:

- instructions on how to use the program
- instructions on how to set up your data-file
- some tips for model specification. For an exhaustive explanation of how to specify a model and examples, please refer to the official manual of mkfm6 (included with the program).
- an error report (found and solved other errors? Please mail these to 5664594@student.uva.nl so I can add them to this how-to).

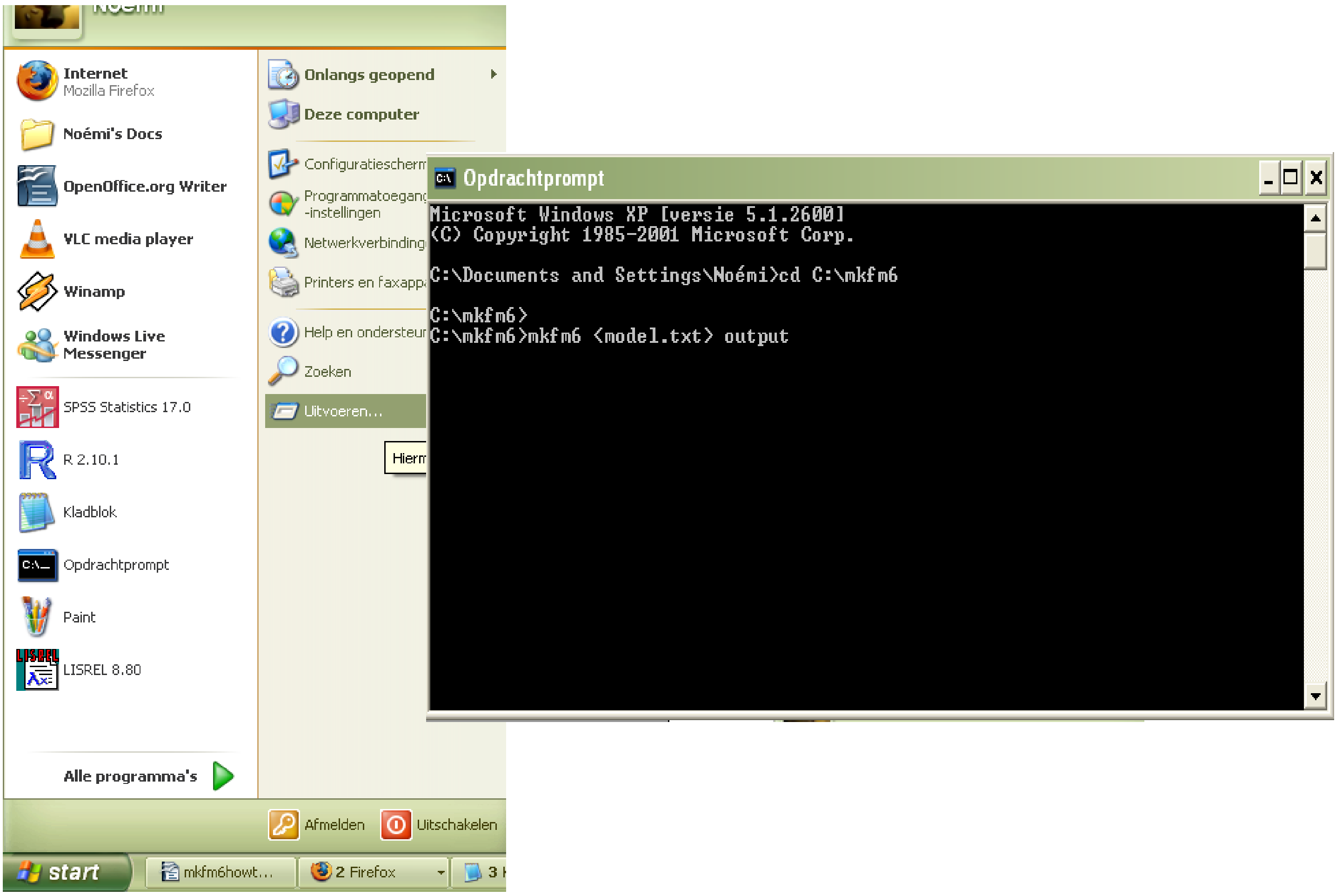
How to use the program (reading data and producing output)

This program works using a prompt. First you need to open a prompt. In windows you can find under the start button the button “run”(dutch:Uitvoeren). Click this button.
You now get a prompt window. Set your working directory (with all your data and input files and mkfm6 in it). You do this by typing “ cd directory “ and press enter.

Example: My directory of choice is C:\mkfm6. First I created this directory myself and then put all my mkfm6 files in it. Then I type in: cd C:\mkfm6 and then press enter. Now this directory is set as my working directory.

To run the analysis, type in “ mkfm6 <inputfile> outputfile and press enter.

Example: my inputfile is called “model.txt” and I'll call my outputfile “output”.
So I type in mkfm6 <model.txt> output and press enter.



Now the outputfile “output” will appear in your working directory. Open it (with notepad) to see the results of your analysis.

How to prepare your datafile

Before you start analyzing your data you should of course make sure your data is presented in a way so so mkfm6 can deal with it. Firstly, your datafile should be in the form of any kind of plain text (.txt extension is fine). On the first line, specify the number of time points. On the second line, you need to specify a[0]. You can fill in the means of the observed (y) variables here, or zeros for instance. It should not make a big difference for the results.

Next follows the actual data. The first row should contain the data on time point 1 for each y-variable and fixed regressors. The second row has the data for time point two, etc. Each column should represent one y-variable or fixed regressor. The y-variables should always be first, the fixed regressors should follow after. So, a dataset with two y-variables, one fixed regressor, and ten time points could look like this:

```
10
3.6 4
1	3	1
2	1	1
4	4	1
5	5	1
7	7	1
4	4	2
5	5	2
3	6	2
3	2	2
2	3	2
```

Some important notes:

- Leave one empty line after the data (because of a possible bug in mkfm6).
- Also, make sure you have not used commas instead of dots for decimal points. One easy way to rectify this is to choose 'replace all' in notepad and replace the commas with dots.
- Y-variables can have missings, but the fixed regressors cannot have missings.
- When you save/name your datafile, make sure you do not use more than 12 characters, including the file extension.
- If you use tab delimited data, do not let the data from one variable spill over into the other, this could lead to some funny results.

So for instance not:

```
1	3	1
2	1	1
4	4.33333		1	← None of this
5	5.33333		1	← None of this
7	7	1
4	4	2
5	5	2
3	6	2
3	2	2
```

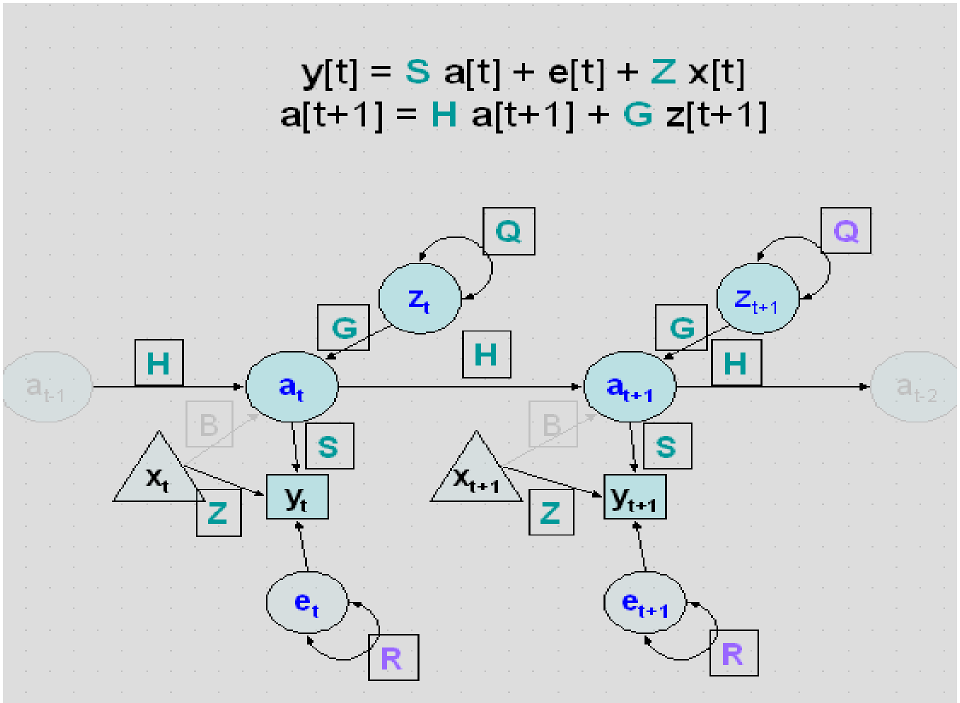
How to specify your model

With mkfm6, it is possible to model autoregressive (AR), moving-average (MA), and autoregressive moving average (ARMA) models, with multiple orders (lags), indicators (vector AR models), and subjects. Fixed regressors can be modeled, as well as latent variables.

I will not get too deep into the issue of model specification since it has been excellently documented in the official manual of mkfm6, including examples of several possible models. I will include here a path model I "borrowed" from a lecture by Dr. C. Dolan (next page). This is nice to have, in order to visualize the model you are specifying, and to see from the image which vectors you need to specify in your model file. It is a good idea to draw your model before you start to specify it.

Some tips and notes:

- When you use latent variables, do not forget to scale them in your model.
- When you are not using latent variables, don't forget you do *not* need to specify scaling (especially for LISREL users).
- Don't forget to specify P. I set them to the square root of the mean's standard error (Variance/N), but you can try different values (it should not matter too much for the results).
- If you do not want to use latent variables, you fix S to ones (the same amount of latent variables as y-variables) and R to zero. Now the "latent variables" have a one-on-one relationship with the y-variables (they are the same).
- You might wonder how to model multiple orders. This can be done in the following way, for instance with an order 2 model with one indicator: Specify one-on-one relationships between three latent variables and the one indicator (S fi –> 1 1 1). Then model the order relationships in H (H fr → 0 0 0   2 0 0   0 3 0). See the mkfm6 manual page 15 and 16 for details, output interpretation, and an alternative specification.
- Leave one empty line at the end of your model file.



ny = number of y-variables

ne = number of latent variables

The following are the random variables in the model (taken from mkfm6 manual):

y[t] observed random variable (NY x 1)

a[t] latent random variable (NE x 1)

e[t] latent (residual) random (NY x 1)

z[t] latent (residual) random (NE x 1)

x[t] observed fixed exogenous regressors (NX x 1)

These are the parameter matrices from the model that are used for the model specification (taken from the mkfm6 manual):

S (NY x NE matrix) regression parameters **y**[t] on **a**[t] (LISREL: Λ_y)

d (NY x 1 vector) intercept in regression of **y**[t] on **a**[t] (LISREL: τ_y)

H (NE x NE matrix) regression parameters **a**[t+1] on **a**[t] (LISREL: **B**)

c (NE x 1 vector) intercept in regression of **a**[t+1] on **a**[t] (LISREL: α)

R (NY x NY symmetric matrix) covariance matrix of **e**[t] (LISREL: Θ_ε)

Q (NE x NE symmetric matrix) covariance matrix of **z**[t] (LISREL: Ψ)

Z (NY x NX matrix) regression parameters **y**[t] on **x**[t]

G (NE x NE matrix) regression parameters **a**[t] on **z**[t]

How to read the output

The output of mkfm6 is quite straightforward. However, one thing that may be good to mention is what the direction of the arrows should be when interpreting the H matrix.

H fr parameters

	$Y1_t$	$Y2_t$	$Y3_t$
$Y1_{t+1}$	1	2	3
$Y2_{t+1}$	4	5	6
$Y3_{t+1}$	7	8	9

So:

1: $Y1_t \rightarrow Y1_{t+1}$
2: $Y2_t \rightarrow Y1_{t+1}$
3: $Y3_t \rightarrow Y1_{t+1}$
4: $Y1_t \rightarrow Y2_{t+1}$
5: $Y2_t \rightarrow Y2_{t+1}$
6: $Y3_t \rightarrow Y2_{t+1}$
7: $Y1_t \rightarrow Y3_{t+1}$
8: $Y2_t \rightarrow Y3_{t+1}$
9: $Y3_t \rightarrow Y3_{t+1}$

How to deal with errors

Message in output: "fatal...error number 6200
message: datafile not found "

While you are sure you

1 – Specified the name correctly in your model input file
2 – Placed the datafile in the correct folder

Then the filename of the datafile is probably too long. Shorten it and try again.

Message in output: "fatal...error number 5225
message: nx and Z= inconsistent" (This is also applicable for ne or ny in combination with H and S.)

The number of nx you filled in is not consistent with the Z you gave. Maybe your Z-matrix is too large or too small, or you have specified nx=1 (or higher) while Z=0.

Example:
mo=1 ny=3 ne=3 **nx=4**
df=datcase121 rf= ns=1 mi=999
S=1 H=1 Q=1 d=1 c=0 **Z=0** G=1 R=0 P=1

Message in output: "fatal...error number 6020
message: expecting "fr" or "fi", but not found "

Maybe you specified only the fr (free) or (fi) parts for one or more matrices. Specify both the fr *and* fi part for each matrix you specified to be fixed/estimated (LISREL users may forget to do this, as it is not necessary in LISREL).

Example:
S=1 H=1 Q=1

S fi di
1 1 1
S fr di
0 0 0

H fi
0 0 0
0 0 0
0 0 0
H fr
2 3 4
5 6 7
8 9 10

! Q fi is missing here!

Q fr di
43 44 45

Message in output: “fatal...error number 6025
message: matrix name incorrect (not S,R,Q,H,P,d,c,Z,G) “
This one is pretty clear, you probably specified a matrix that does not exist in mkfm6. Check for typos.

Example:
nm=1 se=yes
mo=1 ny=3 ne=3 nx=4
df=dat121.dat rf=no ns=1 mi=999
S=1 H=1 Q=1 d=1 c=0 Z=1 **J=1** G=1 R=0 P=1

Message in output (Returns estimations, but no standarderrors etc. Produces a warning):
“warning: error st.err   5400”
This probably means there is something wrong with your starting values, but could be anything that causes the standard errors to not be estimated. Try counting if you have the right number of starting values and check if you have any strange starting values specified.