Temporal stability of symptoms of affective disorders, cognitive vulnerability and personality over time

Struijs, S.Y.; Lamers, F.; Verdam, M.G.E.; Ballegooijen, W. van; Spinhoven, P.; Does, W. van der; Penninx, B.W.J.H.

Journal of Affective Disorders 260 (2020) 77–83

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Research paper

Temporal stability of symptoms of affective disorders, cognitive vulnerability and personality over time

Sascha Y. Struijs[a,b,*], Femke Lamers[b], Mathilde G.E. Verdam[a], Wouter van Ballegooijen[b,c], Philip Spinhoven[a,d], Willem van der Does[a,d], Brenda W.J.H. Penninx[b]

[a] Institute of Psychology, Leiden University, AK Leiden 2333, the Netherlands
[b] Amsterdam UMC, Vrije Universiteit Amsterdam, Department of Psychiatry, Amsterdam Public Health, Amsterdam 1081 HV, the Netherlands
[c] Section Clinical Psychology, Vrije Universiteit Amsterdam, Amsterdam 1081 BT, the Netherlands
[d] Department of Psychiatry, Leiden University Medical Centre, 2333 ZA, the Netherlands

ARTICLE INFO

Keywords:
Temporal stability
Affective disorder
Cognitive vulnerability
Depressive disorder
Anxiety disorder
Personality

ABSTRACT

Background: Signs and symptoms of psychopathology can be chronic but are generally regarded as less stable over time than markers of cognitive vulnerability and personality. Some findings suggest that these differences in temporal stability are modest in size but a rigorous examination across concepts is lacking. The current study investigated the temporal stability of affective symptoms, cognitive vulnerability markers and personality traits at various assessments over nine years.

Methods: Participants of the Netherlands Study of Depression and Anxiety were assessed at baseline and re-assessed after 2, 4, 6 and 9 years. They were grouped on the basis of waves of depression and anxiety CIDI-diagnoses into stable healthy (n = 768), stable patients (n = 352) and unstable patients (n = 821). We determined temporal stability by calculating intraclass correlation coefficients (ICC) and consistency indices of latent state-trait analyses (LST).

Results: Temporal stability was moderate to high for symptoms (range ICC's 0.54–0.73; range consistency 0.64–0.74), cognitive vulnerability (range ICC's 0.53–0.76; range consistency 0.60–0.74) and personality (range ICC's 0.57–0.80; range consistency.60 -0.75). Consistency indices for all measures were on average a bit lower in the unstable group (ICC = 0.54) compared to the stable groups (ICC = 0.61). Overall stability was similarly high after 2, 4, 6 and 9 years.

Conclusion: The 9-year stability over time of symptoms of affective disorders and that of indices of cognitive vulnerability and personality are remarkably similar and relatively high.

1. Introduction

A general consensus exists that mental disorders and symptoms of psychopathology are transient states, while personality dimensions are traits. In between personality and symptoms are cognitive vulnerabilities, such as biases in attention or dysfunctional thoughts (Hong and Cheung, 2015; Jeronimus et al., 2016). This implies that symptoms and disorders wax and wane over time, whereas personality traits remain relatively stable. Cognitive vulnerabilities are thought to take an intermediate position.

Affective disorders are defined as episodic disorders, but they may become chronic (Spijker et al., 2002). Recent research on the course of these disorders has revealed that chronicity is more the rule than the exception (Kessing and Andersen, 2017; Verduijn et al., 2017). For example, over 60% of primary care patients with major depressive disorders had stable or high residual symptoms (Verhoeven et al., 2018). Similar results were found in anxiety disorders, with about 60% of participants showing an unfavourable course (Spinhoven et al., 2016a). Test-retest correlations of depression and anxiety symptoms range from 0.30 to 0.70 with lower scores for longer time-intervals (Ormel et al., 2013).

Research on the temporal stability of personality generally supports the view that personality traits are relatively stable. The temporal stability of personality was moderate to high in adult samples (Ferguson, 2010; Roberts and DelVecchio, 2000). The courses of locus of control, positive affect and negative affect were also relatively stable over a 9-year time interval (Wight et al., 2003). Depressive symptoms, anxiety sensitivity and behavioural inhibition had moderate test-retest

* Corresponding author at: Institute of Psychology, Leiden University, AK Leiden 2333, the Netherlands.
E-mail address: s.y.struijs@fsw.leidenuniv.nl (S.Y. Struijs).

https://doi.org/10.1016/j.jad.2019.08.090
Received 25 June 2019; Received in revised form 22 August 2019; Accepted 28 August 2019
Available online 29 August 2019
0165-0327/

correlations in 606 outpatients for all constructs (Rosellini et al., 2011). And a study examining 266 adolescents from the general population over 10 years on neuroticism and general psychopathology reported similarly high stability over time for both constructs (Aldinger et al., 2014). A meta-analysis combing 984 test-retest correlations over periods of 1 to 14 years concluded that the majority of reliable variance in personality traits is attributable to stable influences (83%), which is higher than the attributable variance of affect, self-esteem and life-satisfaction (42 to 56%) (Anusic and Schimmack, 2016). The difference in stability is modest, however, and is about 33% higher for neuroticism than for symptoms of affective disorders (Ormel et al., 2013).

Overall, studies seem to converge on the conclusion that the temporal stability of all measures is moderate to high, and that personality measures are indeed somewhat more stable over time than the other measures. The temporal stability of symptoms, cognitive vulnerability and personality traits has never been examined in combination in a large cohort. Recent methodological advances have provided a new framework (the latent state-trait theory (Steyer et al., 2015) to examine the temporal stability of various constructs.

The present study aims to examine the temporal stability of various measures of symptoms of affective disorders, cognitive vulnerability and personality using multiple methods. We analysed data from the Netherlands Study of Depression and Anxiety (NESDA) cohort. Symptom measures of affective disorders are operationalized by means of commonly used self-report questionnaires for symptoms of depression, anxiety and fear. The specific measures of cognitive vulnerability and personality are operationalized by known risk factors for affective disorders: rumination, hopelessness and worry for cognitive vulnerability; neuroticism, extraversion and locus of control for personality. We expected that the temporal stability will be strongest for personality traits, weakest (but still moderately high) for symptoms, and intermediate for indices of cognitive vulnerability. Both short-term and longer-term temporal stability were assessed. We explored how stable all the constructs are in patients who lose or gain their diagnosis over time.

2. Method

2.1. Participants and procedures

Data were derived from the baseline, 2-, 4-, 6- and 9-year follow-up assessments of The Netherlands Study of Depression and Anxiety (NESDA), a longitudinal cohort study of the course and consequences of depression and anxiety. Participants were recruited from different settings: the community, primary care and mental health care. Inclusion criteria were: age 18 through 65; proficiency in the Dutch language; no diagnosis of a psychotic disorder, obsessive compulsive disorder, bipolar disorder or severe addiction disorder. The study protocol was approved by the ethical review board of each participating centre. All participants signed written informed consent before participating in the study. The assessments were conducted from 2004 until 2017. A detailed description of the NESDA design and sampling procedure is provided elsewhere (Penninx et al., 2008). The baseline assessment (T0) was completed by 2981 participants, of whom 2596 (87%) completed the 2-year follow-up (FU2) assessment, 2402 (81%) completed the 4-year follow-up (FU4) assessment, 2256 (76%) completed the 6-year follow-up (FU6) assessment, and 2069 (69%) completed the 9-year follow-up (FU9) assessment. DSM-IV depressive (Major Depressive Disorder) and anxiety disorders (Panic Disorder, Social Anxiety Disorder, Generalized Anxiety Disorder and Agoraphobia) were assessed by means of the Composite International Diagnostic Interview (Wittchen, 1994).

3. Measures

3.1. Symptoms of affective disorders

The 30-item Inventory of Depressive Symptomatology Self Report (Rush et al., 1996) was used to assess severity of depressive symptoms. The IDS-SR has demonstrated satisfactory psychometric qualities with good internal consistency, good convergent validity and high sensitivity to change (Rush et al., 1986; Trivedi et al., 2004). The Cronbach's α for the IDS-SR scale at baseline was 0.92

The 21-item self-report Beck Anxiety Inventory (Beck et al., 1988) was used to measure the severity of anxiety symptoms. The reliability and validity of the BAI are good (Beck et al., 1988; Osman et al., 2002). The Cronbach's α for the BAI at baseline was 0.93.

The Fear Questionnaire (FQ) was used to assess external avoidance behaviour. The FQ is a reliable and valid measure (Zuuren, 1988), and showed high internal consistency at baseline (Cronbach's $\alpha = 0.88$).

3.2. Cognitive vulnerability measures

Rumination and hopelessness were assessed using the Rumination on Sadness and Hopelessness Reactivity subscales of the Leiden Index of Depression Sensitivity-Revised (Van der Does, 2002; Solis et al., 2017). When completing the LEIDS-R, participants are asked to indicate to what extent their thinking changes when they experience mild dysphoria. In the present sample, internal consistencies of the subscales were good at baseline; $\alpha = 0.82$ (LEIDS-RUM) and $\alpha = 0.86$ (LEIDS-HOP).

The Penn State Worry Questionnaire (Meyer et al., 1990) is a 16-item inventory that assesses the general tendency to engage in excessive worry (e.g. 'Once I start worrying, I cannot stop'). In NESDA, the abbreviated 11-item version was used (only Worry Engagement) as this version was proven psychometrically most sound (Fresco et al., 2002). The Cronbach's α for the PSWQ scale in the current study at baseline was 0.96

3.3. Personality measures

We used the subscales Neuroticism (NE) and Extraversion (EX) of the Dutch NEO five-factor inventory (Hoekstra et al., 1996). High NE indicates a propensity to experience negative emotions and low EX indicates a tendency to behave in a reserved and solitary fashion. The internal consistencies of the subscales in the present sample at baseline were good; $\alpha = 0.90$ (NE) and $\alpha = 0.84$ (EX).

The Pearlin Mastery Scale (Pearlin and Schooler, 1978) was used to measure locus of control, i.e., the degree to which individuals believe that they have control over outcomes in their lives. Internal consistency at baseline was good ($\alpha = 0.87$).

3.4. Data analysis

Two-Way random ICCs were calculated for every construct we measured, which indicate the absolute agreement between variables measured at two time points within subjects (McGraw and Wong, 1996). We compared ICCs over 2,4,6 and 9 years. In order to be included in these analyses, participants needed to have complete data on all constructs for the corresponding assessment waves. Because the NEO-FFI was not included at the 6- and 9-year assessment waves, no ICCs could be calculated for these waves.

To further investigate the temporal stability of all constructs, we applied the trait-state model (TSM) proposed by Hamaker et al. (2017) within the LST-R framework. In preparation of these analyses we created two parallel item 'parcels' of all constructs (Little et al., 2013). The item parcels were created based on item-total correlations (ITC) of all items for every questionnaire. The items were assigned to the parcels in such a fashion that the average ITC per parcel was approximately equal.

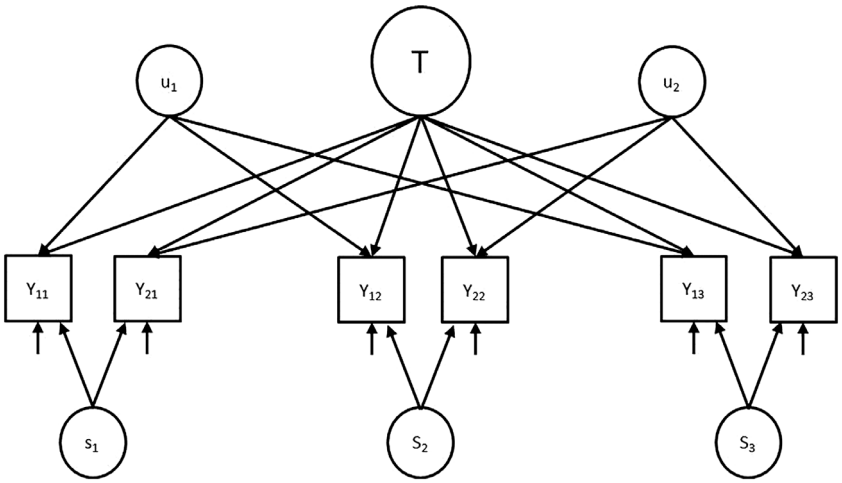

Fig. 1. Model composing observed variances into trait and state components. Y_{ij} indicates the *i*th item parcel for the *j*th assessment wave. T indicates the common trait, u_i indicates the unique trait of the *i*th item parcel, S_j indicates the common state for the *j*th assessment wave.

Application of the TSM to our data allows us to distinguish between time-invariant trait-like sources of variance and time-varying state-like sources of variances in the observations. The TSM was fitted to the means and (co)variances of the observed variables of each construct of interest from at least three waves of measurement using the open source Lavaan package in R (Rosseel, 2012). We choose to first fit the model with data from the baseline assessment wave and the 2- and 4-year follow-up assessment waves, because by doing so, we could analyse all 9 constructs of interest (i.e. neuroticism and extraversion could be analysed in this model) and compare the results. Second, we fitted the model with data from all assessment waves so we could compare the results over different time periods (i.e. 4 years versus 9 years).

The first TSM model was specified such that there is (a) one common trait (T) that directly influences both parcels at all three measurement occasions; (b) two unique traits (u) that directly influence the same parcel at different occasions; (c) one common state (S) per occasion (i.e., three in total), which directly influence the parcels within the same occasion; and (d) measurement error (see Fig. 1 for a graphical representation of this model). As the TSM was fitted to two observed variables at each measurement occasion, additional restrictions were needed to identify the model (Hamaker et al., 2017). Following examples by Zimmermann et al. (2017) and Geiser et al. (2016) we chose to constrain all factor loadings to 1. This implies that each parcel is equally indicative of the underlying common trait and state factors, and that the unique trait variances are invariant over time. The parameters of all models were estimated by the maximum likelihood procedure. To evaluate model fit we report the Chi-squared test of fit, the Comparative Fit Index (CFI), the Tucker Lewis Index (TLI) and the Root Mean Square Error of Approximation (Schermelleh-Engel et al., 2003). The CFI and TLI should preferably be >0.95 and the RMSEA should not be larger than 0.10 for models to provide adequate fit. Based on the parameters of the model we computed reliability, consistency and occasion-specificity coefficients for every construct (Steyer et al., 2015). The reliability is the ratio of the amount of explained variance by the latent trait (T) and state variables (S) to the total variance. The consistency is the ratio of the amount of explained variance by the latent trait (T) to the total variance. The occasion-specificity is the ratio of the amount of explained variance by the latent state variables (S) to the total variance.

In order to be included in the first TSM analysis, participants needed to have complete data on all constructs on the baseline, 2- and 4- year follow-up assessments. This resulted in a final sample of 1941 participants. We checked whether the consistency of constructs would be influenced by the course of CIDI-diagnoses by repeating the analyses while including a grouping variable based on course of diagnoses: a stable diagnosed group that had at least one CIDI diagnosis at every assessment wave ($n = 352$), a stable undiagnosed group that had no CIDI diagnosis at any assessment wave ($n = 768$) and a unstable group of all participants who had at least one CIDI diagnosis at any assessment wave but not every assessment wave ($n = 821$).

We also checked whether the consistency of constructs would be influenced by the length of the follow-up period. This was done by repeating the analyses in a model that also included data from the 6- and 9-year follow up assessment waves in the full sample. In order to be included in these long-term state-trait analyses, participants needed to have complete data on all constructs on all assessments. This resulted in a final sample of 1510 participants.

4. Results

4.1. Descriptive findings

Table 1 shows sociodemographics, clinical characteristics and psychological vulnerabilities of the sample. The majority of the sample consisted of women and the average age at baseline was 42 years. The percentage of participants with a diagnosis of a depressive of anxiety disorder decreases over time from 57% to 27%. There were few missing values for all constructs on all assessment waves which ranged from 1% to 6%. At baseline drop-outs (participants that missed one or more assessment waves, $n = 1096$) were similar to completers (participants that completed all assessment waves, $n = 1885$) in age and sex but significantly differed for all other measured variables with drop-outs having more diagnoses and higher scores than completers. These differences were small to moderate however (Cohen's d ranging from 0.13 to 0.45 and $\varphi = 0.17$).

4.2. Intraclass correlation coefficients

Table 2 shows the ICCs of all constructs calculated from baseline to the 2, 4, 6 and 9-year follow-up assessments. ICCs ranged from moderate to strong for all constructs. On average the ICCs tended to slightly decrease over time; this pattern was visible for all constructs. The only non-overlapping confidence intervals between constructs per assessment wave were that of EX with all other measures except NE and WOR. This possibly indicates that EX is more stable over time than other measures, however we did not directly test this difference with a significant test. All in all, absolute differences between ICC's over all waves and all constructs were very small, with the largest difference being that of EX and BAI, HOP and RUM at the 4-year interval which is 22% (0.17/0.77).

Table 1
Overview of sociodemographics, psychological symptoms, cognitive vulnerabilities and personality per assessment.

Sociodemographics	Baseline (n = 2981[a]) Mean	sd	2-year (n = 2596[a]) Mean	sd	4-year (n = 2402[a]) Mean	sd	6-year (n = 2256[a]) Mean	sd	9-year (n = 2069[a]) Mean	sd
Age	41.9	13.1								
Years of education	12.2	3.3								
Female (n,%)	1979	66								
Diagnosed with depressive or anxiety disorder (%)	1688	57	952	37	751	31	617	27	553	27
Psychological symptoms										
Depression (IDS)	21.5	14.1	15.8	12.0	15.5	12.0	15.1	11.9	14.8	11.7
Anxiety (BAI)	12.1	10.7	8.7	8.7	8.1	8.5	8.4	8.5	7.7	8.4
Fear (FQ)	24.8	19.9	19.4	17.7	18.2	17.8	17.4	17.2	16.1	17.2
Cognitive vulnerabilities										
Hopelessness (LEIDS-HOP)	4.7	4.6	3.9	4.2	3.6	3.8	3.4	3.8	3.1	3.6
Rumination (LEIDS-RUM)	9.0	5.2	8.0	5.2	7.2	4.8	6.7	4.7	6.3	4.7
Worry (PSWQ)	30.8	11.9	28.0	11.7	25.9	11.3	26.1	11.5	25.8	11.4
Personality										
Neuroticism (NEO-FFI)	36.3	9.4	33.5	9.0	32.7	8.5	–	–	–	–
Extraversion (NEO-FFI)	36.9	7.4	37.8	7.2	38.1	7.4	–	–	–	–
Locus of control (PM)	17.3	4.6	18.3	4.5	18.8	4.7	18.8	4.6	19.0	4.7

Note. IDS = Inventory of Depressive Symptomatology, BAI = Beck Anxiety Inventory, FQ = Fear Questionnaire, LEIDS-HOP = Leiden Index of Depression Sensitivity-Hopelessness subscale, LEIDS-RUM = Leiden Index of Depression Sensitivity-Rumination subscale, NEO-FFI = NEO Five-Factor Inventory, PM = Pearlin Mastery Scale.

[a] This number indicates the maximum number of participants for each assessment wave. There are missing values and the number of these missing values differ per construct and assessment wave. This is described in detail in the methods section. - = not assessed.

4.3. Consistency coefficients from the TSM model

Table 3 shows the reliability, occasion specificity, consistency and model fit for each of the constructs in the full sample over the baseline, 2-year follow-up and the 4-year follow-up assessments (n = 1941). Model fit was acceptable for all constructs. The reliability of all constructs was good (ranging from 0.83 to 0.93). On average the consistency of all constructs was 0.69, while the occasion-specificity was 0.19. This means that on average the ratio of explained variance by the latent trait factors on the total explained variance of the models was 78%, indicating strong temporal stability of all constructs. There were some differences between consistency coefficients that ranged from 0.62 to 0.75. However, these differences occurred within each domain of the constructs (psychological symptoms, psychological vulnerabilities and personality) and remained relatively small. Results from analyses in which we used a Full Information Maximum Likelihood (FIML) procedure which handles missing values are largely identical (see supplemental Table 2).

Supplemental Table 1 shows the reliability, occasion specificity, consistency and model fit for each of the constructs in participants with no diagnosis at any assessment (stable healthy), diagnoses at every assessment (stable diagnosis) or participants with an alternating presence of a diagnosis among assessments (unstable) over the baseline, 2-year follow-up and the 4-year follow-up assessments (n = 1941). Model fit was acceptable for all constructs. The reliability of all constructs was good (ranging from 0.69 to 0.91). On average the consistency of all constructs was 0.60 in the stable healthy group, 0.61 in the stable diagnosis group and 0.54 in the unstable group. The consistency was significantly stronger ($p < .003$) for the stable healthy versus the unstable group for IDS, WORRY and NE. The consistency was significantly stronger ($p < .003$) for the stable diagnosis versus the unstable group for IDS and BAI. The latent trait factor for IDS and BAI in the unstable group still explained the most variance relative to the total explained variance for those constructs: 53% and 52% respectively.

Reliability, occasion specificity, consistency and model fit for each of the constructs in the full sample over the baseline, 2-year follow-up, 4-year follow-up, 6-year follow-up and 9-year follow-up assessments (n = 1520) are provided in supplemental Table 3. Model fit was acceptable for all constructs. The average consistency of all constructs (excluding NE and EX) was 0.66 while the occasion-specificity was 0.21. Overall, results were very similar to the results of the TSM model over the baseline, 2-year follow-up and the 4-year follow-up

Table 2
Intraclass correlation coefficients (ICC) calculated from baseline to the 2, 4, 6 and 9-year follow-up assessments with a 95% Confidence Interval in parentheses.

	T0-FU2 ICC (n = 2154)	T0-FU4 ICC (n = 2119)	T0-FU6 ICC (n = 1953)	T0-FU9 ICC (n = 1786)
Psychological symptoms				
Depression (IDS)	.67 (0.52–0.76)	.63 (0.50–0.72)	.62 (0.49–0.70)	.59 (0.47–0.68)
Anxiety (BAI)	.67 (0.59–0.73)	.60 (0.51–0.67)	.58 (0.52–0.63)	.54 (0.45–0.62)
Fear (FQ)	.73 (0.66–0.78)	.68 (0.60–0.74)	.67 (0.57–0.73)	.63 (0.49–0.73)
Cognitive vulnerabilities				
Hopelessness (LEIDS-HOP)	.70 (0.67–0.73)	.60 (0.54–0.65)	.57 (0.50–0.62)	.57 (0.46–0.65)
Rumination (LEIDS-RUM)	.72 (0.68–0.75)	.60 (0.48–0.69)	.57 (0.40–0.68)	.53 (0.32–0.67)
Worry (PSWQ)	.76 (0.71–0.80)	.65 (0.50–0.75)	.65 (0.53–0.73)	.62 (0.49–0.71)
Personality				
Neuroticism (NEO-FFI)	.76 (0.69–0.81)	.70 (0.58–0.78)	–	–
Extraversion (NEO-FFI)	.80 (0.78–0.82)	.77 (0.75–0.79)	–	–
Locus of control (PM)	.67 (0.63–0.71)	.61 (0.53–0.68)	.58 (0.50–0.63)	.57 (0.48–0.63)

Note. IDS = Inventory of Depressive Symptomatology, BAI = Beck Anxiety Inventory, FQ = Fear Questionnaire, LEIDS-HOP = Leiden Index of Depression Sensitivity-Hopelessness subscale, LEIDS-RUM = Leiden Index of Depression Sensitivity-Rumination subscale, NEO-FFI = NEO Five-Factor Inventory, PM = Pearlin Mastery Scale.
- = not assessed.

Table 3
Reliability, occasion specificity, consistency and model fit for each of the constructs in the full sample over the baseline, 2-year follow-up and the 4-year follow-up assessments (n = 1941).

	Rel	Spe	Con (CI)	Con/Rel	χ^2	CFI	TLI	RMSEA
Construct								
Psychological symptoms Depression (IDS)	.90	.21	.69 (0.67–0.71)	.77	21.62	.99	.99	.03
Anxiety (BAI)	.89	.25	.64 (0.61–0.67)	.72	172.15	.99	.97	.10
Fear (FQ)	.91	.17	.74 (0.72–0.76)	.81	20.01	.99	.99	.03
Cognitive vulnerabilities								
Hopelessness (LEIDS-HOP)	.81	.19	.62 (0.60–0.65)	.77	140.25	.98	.97	.09
Rumination (LEIDS-RUM)	.83	.19	.64 (0.63–0.67)	.77	107.91	.99	.98	.08
Worry (PSWQ)	.93	.20	.74 (0.72–0.76)	.80	89.20	.99	.99	.07
Personality								
Neuroticism (NEO-FFI)	.89	.14	.75 (0.73–0.76)	.84	14.60	1.00	.99	.02
Extraversion (NEO-FFI)	.85	.10	.75 (0.73–0.76)	.89	9.35	1.00	1.00	.01
Locus of control (PM)	.85	.22	.63 (0.61–0.65)	.74	12.59	1.00	1.00	.01

Note. IDS = Inventory of Depressive Symptomatology, BAI = Beck Anxiety Inventory, FQ = Fear Questionnaire, LEIDS-HOP = Leiden Index of Depression Sensitivity-Hopelessness subscale, LEIDS-RUM = Leiden Index of Depression Sensitivity-Rumination subscale, NEO-FFI = NEO Five-Factor Inventory, PM = Pearlin Mastery Scale, CFI = Comparative Fit Index, TLI = Tucker–Lewis Index, RMSEA = Root Mean Square Error of Approximation. All models have 9 degrees of freedom and the p-values for the χ^2-tests of model fit were smaller than 0.01 for all models except for Fear (0.02) and locus of control (0.18).

assessments (Table 3), indicating that there is little difference in consistency coefficients when longer time intervals are used for the assessments.

5. Discussion

We examined the temporal stability of symptoms of affective disorders, cognitive vulnerability and personality across 2-, 4-, 6- and 9 years of follow-up using multiple methods in a large group of participants from the Netherlands Study of Depression and Anxiety (NESDA) cohort. The (long-term) temporal stability across all constructs was remarkably similar and relatively high with a small decay over time. This indicates that the state-trait distinction might be less significant than often thought.

Overall, the results support the hypothesis that the temporal stability of the examined constructs is moderate to strong and diminishes only slightly over time. This is in line with previous studies examining the temporal stability of various concepts such as mental disorders (Spinhoven et al., 2016a; Verhoeven et al., 2018), psychopathology symptoms (Nivard et al., 2015; Ormel et al., 2013), personality (Ferguson, 2010; Nivard et al., 2015; Roberts and DelVecchio, 2000) or a combination of those measures including cognitive vulnerabilities (Aldinger et al., 2014; Anusic and Schimmack, 2016; Association, 2000; Ormel et al., 2013; Prenoveau et al., 2011; Rosellini et al., 2011; Wight et al., 2003). This seems to imply that past a certain age, different kinds of psychological vulnerabilities and corresponding signs and symptoms are relatively set (Caspi and Moffitt, 2018) which is in line with recent study results on the course of affective disorders revealing that chronicity is more the rule than the exception (Kessing and Andersen, 2017; Verduijn et al., 2017) The differences we found in temporal stability between symptoms of affective disorders and personality are smaller than the 33% difference reported previously (Ormel et al., 2013). Across all methods we used, the largest difference we observed was 22%. This indicates that the state-trait distinction might be less significant than often thought. This idea fits well with the finding that heritability estimates of neuroticism and common mental disorders are typically within the same range (Ormel et al., 2013) and the finding that neuroticism and depressive symptoms are substantially genetically correlated (Luciano et al., 2018). It might also be that trait-state associations are dynamic and vary according to the individual's developmental context (Durbin and Hicks, 2014). Persons who struggle early in life experience symptoms of psychopathology that may interfere with their ability to make positive changes in their lives, which may affect personality development. Of course, the developmental cycle can also act in opposite direction as a positive manifold. Both processes end in relatively stable states and traits across the board (Borsboom, 2017).

Previous studies examining temporal stability used homogeneous samples: (recovered) patients or the general population. The current study is the first to compare stability estimates across course groups based on CIDI diagnoses. As the NESDA sample is composed of diagnosed and undiagnosed individuals with different course trajectories, we thought these subgroup analyses would provide extra information regarding the temporal stability in more homogenous subgroups. As expected, the average stability for all constructs was smaller for the unstable group versus the stable groups, however these differences were again marginal. The most pronounced group differences were found for symptoms of depression and anxiety which were less stable in the unstable group. However, in this by definition unstable group, the stable trait factor still explained the most variance over time in depression and anxiety symptoms which by theory constitute the most unstable constructs.

The strengths of the current study include the large sample size, including anxious, depressed and healthy individuals; the long follow-up period and the use of multiple assessment methods of both constructs and temporal stability. A limitation of the study is attrition. However, differences between study drop-outs and study completers were small and results from the FIML analyses were largely identical, which implies that missingness does not meaningfully influence our results. Another limitation is the lack of significance testing of the temporal stabilities between constructs. However, the description of temporal stability estimates along with their confidence intervals provide information that is detailed enough to compare the temporal stabilities. Finally, it has to be noted that the LEIDS-R questionnaire was constructed to measure cognitive vulnerability as risk factor for first onsets of disorders and for relapse in recovered depressed patients and was not originally designed for use in currently depressed individuals. So despite the robust finding of increased LEIDS-R (subscale) scores in depressed individuals (Drost et al., 2012; Hendriks et al., 2014; Spinhoven et al., 2016b; Struijs et al., 2018; Wiersma et al., 2011), we are not certain how this should be interpreted in terms of vulnerability. Scores seem to be independent of residual symptoms however, since LEIDS-RUM scores were associated with scores on the Ruminative Response Scale after adjustment for current symptoms of depression (Moulds et al., 2008).

The stability of symptoms of affective disorders and indices of cognitive vulnerability and personality over time is remarkably similar and relatively high. For many people, the symptoms of affective disorders are relatively stable even over a 9 year period. Pharmacological and psychotherapeutic treatments could therefore focus more explicitly on changing stable traits. This is already possible through the use of

existing therapies targeting state-related psychopathology (Roberts et al., 2017) but may be more effective when treatments are used that are designed specifically towards this goal (Barlow et al., 2017).

Funding/support

The infrastructure for the NESDA study (www.nesda.nl) is funded through the Geestkracht program of the Netherlands Organization for Health Research and Development(Zon-Mw, grant number 10-000-1002) and is supported by participating universities and mental health care organizations (VU University Medical Center, GGZ inGeest, Arkin,Leiden University Medical Center, GGZ Rivierduinen, University Medical Center Groningen, Lentis, GGZ Friesland, GGZ Drenthe, Scientific Institute for Quality of Healthcare (IQ healthcare), Netherlands Institute for Health Services Research and Netherlands Institute of Mental Health and Addiction (Trimbos Institute). BP has received funding - not related to this work - from Jansen Research and Boehringer Ingelheim. FL has received funding from the European Union Seventh Framework Program (FP7/2007–2013) under grant agreement n° PCIG12-GA-2012-334065.

Role of the funding source

The funding sources had no further role in study design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the paper for publication.

Contributors

Author BP designed the study and wrote the protocol. Author SS managed the literature searches and analyses. Author SS undertook the statistical analysis and wrote the first draft of the manuscript. Authors BP, MV, WvB, FL, PS and WVDD revised the paper critically for important intellectual content. All authors contributed to and have approved the final manuscript. All authors declare that they have no conflicts of interest.

Declaration of Competing Interest

All authors declare that they have nothing to disclose.

Acknowledgements

None.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jad.2019.08.090.

References

Aldinger, M., Stopsack, M., Ulrich, I., Appel, K., Reinelt, E., Wolff, S., Grabe, H.J., Lang, S., Barnow, S., 2014. Neuroticism developmental courses - implications for depression, anxiety and everyday emotional experience; a prospective study from adolescence to young adulthood. BMC Psychiatry 14, 210. https://doi.org/10.1186/s12888-014-0210-2.

Anusic, I., Schimmack, U., 2016. Stability and change of personality traits, self-esteem, and well-being: introducing the meta-analytic stability and change model of retest correlations. J. Pers. Soc. Psychol. 110, 766–781. https://doi.org/10.1037/pspp0000066.

Association, A.P., 2000. Diagnostic and Statistical Manual of Mental Disorders, 4th, text ed. Author, Washington, DC.

Barlow, D.H., Farchione, T.J., Bullis, J.R., Gallagher, M.W., Murray-Latin, H., Sauer-Zavala, S., Bentley, K.H., Thompson-Hollands, J., Conklin, L.R., Boswell, J.F., Ametaj, A., Carl, J.R., Boettcher, H.T., Cassiello-Robbins, C., 2017. The unified protocol for transdiagnostic treatment of emotional disorders compared with diagnosis-specific protocols for anxiety disorders a randomized clinical trial. JAMA Psychiatry 74, 875–884. https://doi.org/10.1001/jamapsychiatry.2017.2164.

Beck, A.T., Brown, G., Epstein, N., Steer, R.A., 1988. An inventory for measuring clinical anxiety - psychometric properties. J. Consult. Clin. Psychol. 56, 893–897. https://doi.org/10.1037/0022-006X.56.6.893.

Borsboom, D., 2017. A network theory of mental disorders. World Psychiatry 16, 5–13. https://doi.org/10.1002/wps.20375.

Caspi, A., Moffitt, T.E., 2018. All for one and one for all: mental disorders in one dimension. Am. J. Psychiatry appiajp 201817121383. https://doi.org/10.1176/appi.ajp.2018.17121383.

Drost, J., Van der Does, A.J.W., Antypa, N., Zitman, F.G., Dyck, R, Van, Spinhoven, P., 2012. General, specific and unique cognitive factors involved in anxiety and depressive disorders. Cognit. Ther. Res. 36, 621–633. https://doi.org/10.1007/s10608-011-9401-z.

Durbin, C.E., Hicks, B.M., 2014. Personality and psychopathology: a stagnant field in need of development. Eur. J. Pers. 28, 362–386.

Ferguson, C.J., 2010. A meta-analysis of normal and disordered personality across the life span. J. Pers. Soc. Psychol. 98, 659–667. https://doi.org/10.1037/a0018770.

Fresco, D.M., Frankel, A.N., Mennin, D.S., Turk, C.L., Heimberg, R.G., 2002. Distinct and overlapping features of rumination and worry: the relationship of cognitive production to negative affective states. Cognit. Ther. Res. 26, 179–188.

Geiser, C., Griffin, D., Shiffman, S., 2016. Using multigroup-multiphase latent state-trait models to study treatment-induced changes in intra-individual state variability: an application to smokers' affect. Front. Psychol. 7. https://doi.org/10.3389/fpsyg.2016.01043.

Hamaker, E.L., Schuurman, N.K., Zijlmans, E.A.O., 2017. Using a few snapshots to distinguish mountains from waves: weak factorial invariance in the context of trait-state research. Multivariate Behav. Res. 52, 47–60. https://doi.org/10.1080/00273171.2016.1251299.

Hendriks, S.M., Licht, C.M.M., Spijker, J., Beekman, A.T.F., Hardeveld, F., de Graaf, R., Penninx, B.W.J.H., 2014. Disorder-specific cognitive profiles in major depressive disorder and generalized anxiety disorder. BMC Psychiatry 14, 96. https://doi.org/10.1186/1471-244X-14-96.

Hoekstra, R.A., Ormel, J., Fruyt, F.D., 1996. Handleiding NEO Persoonlijkheidsvragenlijsten. Swets Test Services, Lisse.

Hong, R.Y., Cheung, M.W.-L., 2015. The structure of cognitive vulnerabilities to depression and anxiety: evidence for a common core etiologic process based on a meta-analytic review. Clin. Psychol. Sci. 3, 892–912. https://doi.org/10.1177/2167702614553789.

Jeronimus, B.F., Kotov, R., Riese, H., Ormel, J., 2016. Neuroticism's prospective association with mental disorders halves after adjustment for baseline symptoms and psychiatric history, but the adjusted association hardly decays with time: a meta-analysis on 59 longitudinal/prospective studies with 443 313 pa. Psychol. Med. 46, 2883–2906. https://doi.org/10.1017/S0033291716001653.

Kessing, L.V., Andersen, P.K., 2017. Evidence for clinical progression of unipolar and bipolar disorders. Acta Psychiatr. Scand. 135, 51–64. https://doi.org/10.1111/acps.12667.

Little, T.D., Rhemtulla, M., Gibson, K., Schoemann, A.M., 2013. Why the items versus parcels controversy needn't be one. Psychol. Methods 18, 285–300. https://doi.org/10.1037/a0033266.

Luciano, M., Hagenaars, S.P., Davies, G., Hill, W.D., Clarke, T.-K., Shirali, M., Harris, S.E., Marioni, R.E., Liewald, D.C., Fawns-Ritchie, C., Adams, M.J., Howard, D.M., Lewis, C.M., Gale, C.R., McIntosh, A.M., Deary, I.J., 2018. Association analysis in over 329,000 individuals identifies 116 independent variants influencing neuroticism. Nat. Genet. 50 6-+. https://doi.org/10.1038/s41588-017-0013-8.

McGraw, K.O., Wong, S.P., 1996. Forming inferences about some intraclass correlation coefficients. Psychol. Methods 1, 30–46. https://doi.org/10.1037/1082-989X.1.4.390.

Meyer, T.J., Miller, M.L., Metzger, R.L., Borkovec, T.D., 1990. Development and validation of the penn state worry questionnaire. Behav. Res. Ther. 28, 487–495.

Moulds, M.L., Kandris, E., Williams, A.D., Lang, T., Yap, C., Hoffmeister, K., 2008. An investigation of the relationship between cognitive reactivity and rumination. Behav. Ther. 39, 65–71. https://doi.org/10.1016/j.beth.2007.05.001.

Nivard, M.G., Dolan, C.V., Kendler, K.S., Kan, K.J., Willemsen, G., van Beijsterveldt, C.E., Lindauer, R.J., van Beek, J.H., Geels, L.M., Bartels, M., Middeldorp, C.M., Boomsma, D.I., 2015. Stability in symptoms of anxiety and depression as a function of genotype and environment: a longitudinal twin study from ages 3 to 63 years. Psychol. Med. 45, 1039–1049. https://doi.org/10.1017/S003329171400213X.

Ormel, J., Jeronimus, B.F., Kotov, R., Riese, H., Bos, E.H., Hankin, B., Rosmalen, J.G.M., Oldehinkel, A.J., 2013. Neuroticism and common mental disorders: meaning and utility of a complex relationship. Clin. Psychol. Rev. 33, 686–697. https://doi.org/10.1016/j.cpr.2013.04.003.

Osman, A., Hoffman, J., Barrios, F.X., Kopper, B.A., Breitenstein, J.L., Hahn, S.K., 2002. Factor structure, reliability, and validity of the beck anxiety inventory in adolescent psychiatric inpatients. J. Clin. Psychol. 58, 443–456. https://doi.org/10.1002/jclp.1154.

Pearlin, L.I., Schooler, C., 1978. Structure of coping. J. Health Soc. Behav. 19, 2–21.

Penninx, B.W.J.H., Beekman, A.T.F., Smit, J.H., Zitman, F.G., Nolen, W.A., Spinhoven, P., Cuijpers, P., de Jong, P.J., van Marwijk, H.W.J., Assendelft, W.J.J., van der Meer, K., Verhaak, P., Wensing, M., de Graaf, R., Hoogendijk, W.J., Ormel, J., van Dyck, R., 2008. The netherlands study of depression and anxiety (NESDA): rationale, objectives and methods. Int. J. Methods Psychiatr. Res. 17, 121–140. https://doi.org/10.1002/mpr.256.

Prenoveau, J.M., Craske, M.G., Zinbarg, R.E., Mineka, S., Rose, R.D., Griffith, J.W., 2011. Are anxiety and depression just as stable as personality during late adolescence? results from a three-year longitudinal latent variable study. J. Abnorm. Psychol. 120, 832–843. https://doi.org/10.1037/a0023939.

Roberts, B.W., DelVecchio, W.F., 2000. The rank-order consistency of personality traits from childhood to old age: a quantitative review of longitudinal studies. Psychol. Bull. 126, 3–25. https://doi.org/10.1037/0033-2909.126.1.3.

Roberts, B.W., Luo, J., Briley, D.A., Chow, P.I., Su, R., Hill, P.L., 2017. A systematic review of personality trait change through intervention. Psychol. Bull. 143, 117–U122. https://doi.org/10.1037/bul0000088.

Rosellini, A.J., Fairholme, C.P., Brown, T.A., 2011. The temporal course of anxiety sensitivity in outpatients with anxiety and mood disorders: relationships with behavioral inhibition and depression. J. Anxiety Disord. 25, 615–621. https://doi.org/10.1016/j.janxdis.2011.02.001.

Rosseel, Y., 2012. lavaan: an R package for structural equation modeling. J. Stat. Softw. 48, 1–36.

Rush, A.J., Giles, D.E., Schlesser, M.A., Fulton, C.L., 1986. The inventory for depressive symptomatology (IDS): preliminary findings. Psychiatry Res. 18, 65–87. https://doi.org/10.1016/0165-1781(86)90060-0.

Rush, A.J., Gullion, C.M., Basco, M.R., Jarrett, R.B., 1996. The inventory of depressive symptomatology (IDS): psychometric properties. Psychol. Med. A J. Res. Psychiatry Allied Sci. 26, 477–486. https://doi.org/10.1017/S0033291700035558.

Schermelleh-Engel, K., Moosbrugger, H., Müller, H., 2003. Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. Methods Psychol. Res. 8, 23–74.

Solis, E., Antypa, N., Conijn, J.M., Kelderman, H., Van der Does, W., 2017. Psychometric properties of the Leiden index of depression sensitivity (LEIDS). Psychol. Assess. 29, 158–171. https://doi.org/10.1037/pas0000326.

Spijker, J., Graaf, R.De, Bijl, R.V., Beekman, A.T.F., Ormel, J., Nolen, W.A., 2002. Duration of major depressive episodes in the general population: results from the Netherlands mental health survey and incidence study (NEMESIS). Br. J. Psychiatry 181, 208–213. https://doi.org/10.1192/bjp.181.3.208.

Spinhoven, P., Batelaan, N., Rhebergen, D., van Balkom, A., Schoevers, R., Penninx, B.W., 2016a. Prediction of 6-yr symptom course trajectories of anxiety disorders by diagnostic, clinical and psychological variables. J. Anxiety Disord. 44, 92–101. https://doi.org/10.1016/j.janxdis.2016.10.011.

Spinhoven, P., Drost, J., de Rooij, M., van Hemert, A.M., Penninx, B.W.J.H., 2016b. Is experiential avoidance a mediating, moderating, independent, overlapping, or proxy risk factor in the onset, relapse and maintenance of depressive disorders? Cognit. Ther. Res. 40, 150–163. https://doi.org/10.1007/s10608-015-9747-8.

Steyer, R., Mayer, A., Geiser, C., Cole, D.A., 2015. A theory of states and traits-revised. Annu. Rev. Clin. Psychol. 11 (11) 71-+. https://doi.org/10.1146/annurev-clinpsy-032813-153719.

Struijs, S.Y., Lamers, F., Rinck, M., Roelofs, K., Spinhoven, P., Penninx, B.W.J.H., 2018. The predictive value of approach and avoidance tendencies on the onset and course of depression and anxiety disorders. Depress. Anxiety. https://doi.org/10.1002/da.22760.

Trivedi, M.H., Rush, A.J., Ibrahim, H.M., Carmody, T.J., Biggs, M.M., Suppes, T., Crismon, M.L., Shores-Wilson, K., Toprac, M.G., Dennehy, E.B., Witte, B., Kashner, T.M., 2004. The inventory of depressive symptomatology, clinician rating (IDS-C) and self-report (IDS-SR), and the quick inventory of depressive symptomatology, clinician rating (QIDS-C) and self-report (QIDS-SR) in public sector patients with mood disorders: a psych. Psychol. Med. A J. Res. Psychiatry Allied Sci. 34, 73–82. https://doi.org/10.1017/S0033291703001107.

Van der Does, W., 2002. Cognitive reactivity to sad mood: structure and validity of a new measure. Behav. Res. Ther. 40, 105–119. https://doi.org/10.1016/S0005-7967(00)00111-X.

Verduijn, J., Verhoeven, J.E., Milaneschi, Y., Schoevers, R.A., van Hemert, M., A., Beekman, A.T.F., Penninx, B.W.J.H., 2017. Reconsidering the prognosis of major depressive disorder across diagnostic boundaries: full recovery is the exception rather than the rule. BMC Med. 15, 215. https://doi.org/10.1186/s12916-017-0972-8.

Verhoeven, F.E.A., Wardenaar, K.J., Ruhe, H.G.E., Conradi, H.J., de Jonge, P., 2018. Seeing the signs: using the course of residual depressive symptomatology to predict patterns of relapse and recurrence of major depressive disorder. Depress. Anxiety 35, 148–159. https://doi.org/10.1002/da.22695.

Wiersma, J.E., van Oppen, P., van Schaik, D.J.F., Van der Does, A.J.W., Beekman, A.T.E., Penninx, B.W.J.H., 2011. Psychological characteristics of chronic depression: a longitudinal cohort study. J. Clin. Psychiatry 72, 288–294. https://doi.org/10.4088/JCP.09m05735blu.

Wight, R.G., Aneshensel, C.S., Seeman, M., Seeman, T.E., 2003. Late life cognition among men: a life course perspective on psychosocial experience. Arch. Gerontol. Geriatr. 37, 173–193. https://doi.org/10.1016/S0167-4943(03)00046-3.

Wittchen, H.-.U., 1994. Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI): a critical review. J. Psychiatr. Res. 28, 57–84. https://doi.org/10.1016/0022-3956(94)90036-1.

Zimmermann, J., Mayer, A., Leising, D., Krieger, T., Holtforth, M.G., Pretsch, J., 2017. Exploring occasion specificity in the assessment of DSM-5 maladaptive personality traits a latent state-trait analysis. Eur. J. Psychol. Assess. 33, 47–54. https://doi.org/10.1027/1015-5759/a000271.

Zuuren, F.J.Van, 1988. The fear questionnaire. Some data on validity, reliability and layout. Br. J. Psychiatry 153, 659–662.