# A generic solution for the operationalization and measurement of resilience and resilience processes in longitudinal observations: rationale and basic design of the MARP and LORA studies

**Raffael Kalisch[1,2], Göran Köber[3], Harald Binder[3], Kira Ahrens[4], Ulrike Basten[5], Andrea Chmitorz[1,6,7], Christian J. Fiebach[5,12], Nele Goldbach[4], Rebecca Horstmann[4], Miriam Kampa[1,8,9], Bianca Kollmann[1,6], Klaus Lieb[1,6], Michael Plichta[4], Andreas Reif[4], Anita Schick[2,10], Alexandra Sebastian[6], Michèle Wessa[1,11], Kenneth S.L. Yuen[1,2], Oliver Tüscher[1,6,]*, Haakon Engen[2,]***

[1]Leibniz Institute for Resilience Research (LIR), Mainz, Germany

[2]Neuroimaging Center (NIC), Focus Program Translational Neuroscience (FTN), Johannes Gutenberg University Medical Center, Mainz, Germany

[3]Institute of Medical Biometry and Statistics, Faculty of Medicine, Medical Center University of Freiburg, Freiburg, Germany

[4]Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, University Hospital Frankfurt, Frankfurt (Main), Germany

[5]Department of Psychology, Goethe University, Frankfurt (Main), Germany

[6]Department of Psychiatry and Psychotherapy, Johannes Gutenberg University Medical Center, Mainz, Germany

[7]Faculty of Social Work, Health and Nusing, University of Applied Sciences Esslingen, Esslingen, Germany

[8]Department of Clinical Psychology, University of Siegen, Germany

[9]Bender Institute of Neuroimaging (BION), Department of Psychology, Justus Liebig University, Gießen, Germany

[10]Department of Public Mental Health, Central Institute for Mental Health, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany

[11]Department of Clinical Psychology and Neuropsychology, Institute of Psychology, Johannes Gutenberg University, Mainz, Germany

[12]Brain Imaging Center, Goethe Unviersity Frankfurt, Frankfurt (Main), Germany

*equal contribution


## Abstract

Resilience has been defined as the maintenance or quick recovery of mental health during and after times of adversity. Such good longer-term mental health outcomes despite adversity presumably result from complex and dynamic processes of adaptation to stressor exposure (‘resilience processes’), which in many cases include changes in individual properties. Measuring resilience and identifying resilience processes in observational studies requires longitudinal designs involving repeated and frequent monitoring of mental health, stressor exposure, and potential adaptations. We here present a generic design solution that is currently employed in two cohort studies, the Mainz Resilience Project (MARP) and the Longitudinal Resilience Assessment (LORA). Both projects focus on resilience to everyday life stressors (i.e., microstressors), but we argue that the design scheme is also suitable for studying resilience to macrostressors, or trauma, and can solve some of the pertinent problems of trauma resilience research. We quantify resilience by indexing the reactivity of individuals’ mental health to stressors during a time interval of several months in a ‘stressor reactivity’ (SR) score, derived using a previously introduced residualization approach. SR scores are regularly re-calculated in sliding time windows, to thus build SR time courses that reflect intra-individual temporal variability in resilience. By linking these time courses to repeated measures of (temporally varying) individual properties, resilience processes can be identified. We finish by a discussion of limitations of our approach and potential future developments.

## Main text

The human brain/mind is a complex dynamic system in constant exchange with its internal (bodily) and external environment. External perturbations of the system in the form of stressors may, if minor, be buffered by homeostatic processes that are located within the system itself or involve recruitment of resources from the environment. In this case, there is no lasting change in system function, that is, the coping process leaves the system's mode of operation unaffected. If a perturbation is so strong that it cannot be compensated for by system-environment transactions and exceeds the system's capacity for homeostatic coping, the system will adapt allostatically, by changing its operational set points (Sterling and Eyer, 1988; McEwen BS, 1993; Karatsoreos and McEwen, 2011; Kalisch et al., 2015, 2019). Such allostatic changes can be adaptive, preserving system stability - albeit in a new, different mode of functioning -, or maladaptive, in which case lasting mental health problems will ensue (McEwen BS, 1993).

Resilience, or the preservation of mental health during and after significant stressor exposure (such as traumatizing events, difficult life circumstances, challenging life transitions, or physical illness; Bonanno et al., 2011; Kalisch et al., 2017), then is the result of successful intra-systemic and inter-systemic adaptation processes (Luthar et al., 2000; Sapienza and Masten, 2011; Rutter, 2012; Kent et al., 2014; Bonanno et al., 2015; Kalisch et al., 2017, 2019). The key challenge to resilience research is to identify, describe and, ideally, quantify these adaptation processes. Necessarily, this requires longitudinal investigations (Kalisch et al., 2017).

Longitudinal studies aiming to investigate adaptation processes (also 'resilience processes' in the remainder) must answer five questions. i) What is adversity, or stressor exposure, and how is it measured? ii) How is mental health conceptualized and measured? iii) On the basis of appropriate stressor and mental health assessments: what, exactly, constitutes a resilient outcome? iv) What are the potential adaptation processes and how are they measured? And v) How can the statistical relationship between adaptation processes and resilience as an outcome be established?

In this paper, we will try both to provide general answers to these questions and to describe concrete implementations, as they are currently being employed in two on-going longitudinal resilience studies, Mainz Resilience Project (MARP) and LORA. We consider the basic design scheme of these studies to be of generic nature and hope that it will be useful as a template for other future resilience studies. For the purpose of orientation, we begin by a brief outline of the Mainz Resilience Project (MARP) and Longitudinal Resilience Assessment (LORA) studies.

## 1. The MARP and LORA studies

MARP is conducted by the Leibniz Institute for Resilience Research (LIR; formerly Deutsches Resilienz Zentrum, DRZ) in Mainz, Germany, at its study site at the Neuroimaging Center of the University Medical Center Mainz. MARP started in 2016 and has finished the planned subject recruitment with N=201 in 2019. Subjects are healthy young adults from the greater area of Mainz, aged 18 to 20 years at study inclusion (T0). It is planned that each subject will be followed up over seven years, with the idea of thereby capturing the life phase of transition from a familiar environment (family, school) into academic and/or work life.

This critical life transition is associated with a range of new challenges and in many individuals coincides with the onset or exacerbation of stress-related mental health problems (e.g., Compas et al., 1986; Reavley and Jorm, 2010; Craske et al., 2012; Eisenberg et al., 2012; Herbst et al., 2016; Rotenstein et al., 2016). A further inclusion criterion is the self-reported experience in previous life of at least three adverse events from a canonical list of life events (Canli et al., 2006), based on evidence that cumulative life-time adversity is a risk factor for stress-related psychopathology (Paykel, 2003; Kendler and Gardner, 2010; Seery et al., 2010; Bonanno, 2012; Hammen, 2016). Thus, the MARP sample constitutes an at-risk cohort for the development of symptoms, notably from the anxiety, fear and depression spectrum.

LORA (Longitudinal Resilience Assessment) is conducted by the Collaborative Research Center 1193 ("Neurobiology of Resilience") of the Universities of Mainz and Frankfurt at two study sites (University Medical Center Mainz and Department of Psychiatry, Psychosomatics and Psychotherapy of University Hospital Frankfurt, Germany). LORA started in 2017 and has also finished subject inclusion (T0) at N=1,191 in 2019. LORA subjects are from a wider age range than MARP subjects (18–50 at T0) and have not been selected for previous life events. The cohort therefore does not constitute a particular at-risk sample but is, instead, intended to capture responses to normal life stress in a Western healthy adult population. The expected lower base rate of stress-related symptom development than in MARP is supposed to be compensated for by the comparatively larger sample size. Subjects are recruited from the Rhine-Main metropolitan area, with recruitment hotspots in the cities of Mainz and Frankfurt. The study duration per subject is a minimum of 1.5 years, with further follow-ups depending on the availability of future funding.

Detailed study protocols for MARP and LORA will be published elsewhere. For the purpose of the present report, we briefly introduce the common measurement instruments that serve to longitudinally monitor mental health, stressor exposure and resilience processes in both studies and the temporal sequence in which they are applied.

The main mental health measure used in both MARP and LORA is the 28-item version of the General Health Questionnaire (GHQ-28; Goldberg and Hillier, 1979), an established self-report instrument frequently used by general practitioners or other health professions for the purpose of screening for mental health problems experienced in the past weeks. The GHQ-28 has a four-factor structure and four corresponding sub-scores, each summing over seven items that represent clusters of symptoms or problems in the areas of, respectively, (a) hypervigilance, worrying, and anxiety; (b) psychosocial dysfunction; (c) despair, hopelessness, and depressed mood; and (d) somatic tension, exhaustion, pain, or irritability. Alternatively, the total sum score can be used. The GHQ does not permit clinical diagnoses but can be considered a transdiagnostic instrument (see 2.ii below). In MARP and LORA, subjects fill in the GHQ online every three months (T0, T1, T2, T3, …).

At the same occasion, subjects report on major life events (LE) they experienced in the past three months as well as on daily hassles (DH) they experienced in the last seven days. For LEs, or macrostressors, the studies employ an adapted 27-item version of Canli's 28-item Life Experiences Questionnaire (Canli et al., 2006). Subjects report both the number of occurrences of each event and the degree to which they perceived the event as straining. For DHs, or microstressors, the studies employ a newly generated and validated 58-item instrument (Mainz Inventory of Microstressors, MIMIS) that asks for the number of days out of the last seven days on which a given hassle occurred, as well as for how straining it was (Chmitorz et al., in press).

Together, the regular three-monthly application of the three online instruments establishes a common temporal framework (Fig. 1) that is the backbone of the MARP and LORA studies and serves to establish individual resilience outcomes.

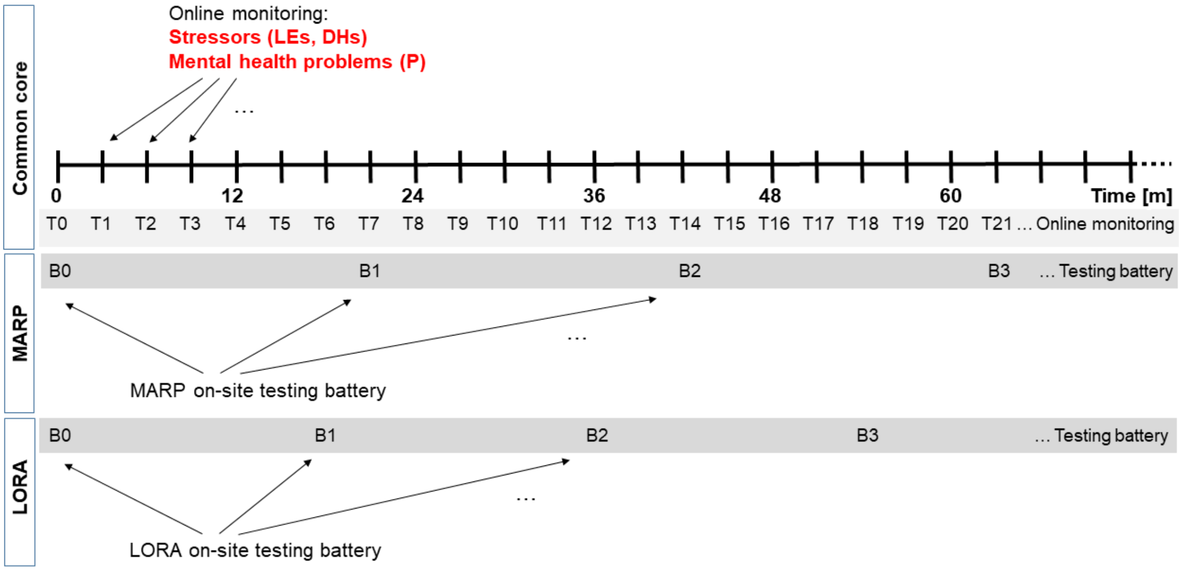



**Figure 1. Generic design scheme employed in MARP and LORA.**

Every three months (T1, T2, …), mental health is assessed via self-report using an online version of the GHQ-28. At the same online monitoring surveys, exposure to macrostressors (life events, LEs) and microstressors (daily hassles, DHs) is reported. Every 1.75 (MARP) and 1.5 (LORA) years (B0, B1, B2, …), subjects visit the study site to complete a study battery.

In order to also monitor potential adaptations in mental, neural and bodily functions, or characteristics - which we assume will permit some subjects to stay mentally healthy or to only show temporary mental health impairments despite stressor exposure - both MARP and LORA employ extensive testing batteries. The batteries are completed by each subject repeatedly in intervals of 1.75 (MARP) and 1.5 years (LORA) at the study site(s) (see B0, B1, B2, … in Fig. 1). Common elements of the battery consist in measures of sociodemographic and socioeconomic variables, lifestyle, personality, coping and appraisal styles, and sampling of blood, hair and stool. In addition, MARP uses a dedicated three-day behavioral, psychophysiological and neuroimaging battery, whereas LORA restricts itself to one day of behavioral and psychophysiological testing. At study inclusion (T0/B0), life-event history and childhood trauma are also assessed.

Detailed descriptions of the battery methods and of the hypotheses guiding the selection of the battery components will be provided elsewhere. In the remainder of the paper, we focus on how the three constituent parts of the MARP and LORA designs (mental health measure, stressor exposure measures, battery) will be combined to investigate resilience processes.

## 2. The conceptual bases of the MARP and LORA designs

### i) What is adversity, or stressor exposure, and how is it measured?

A stressor is a stimulus or situation that elicits a stress response. The immediate practical problem arising from this definition is that what is a stressor for one person may not be a stressor for another person, or also not for the same person in a different phase of his or her life or in a different circumstance. Conversely, a stimulus or situation that is entirely harmless for one may be a significant challenge for another. In the face of this, how can we define an objective stressor or set of stressors to be measured in a resilience study? A further complication is that much of what makes some people resilient to developing mental health problems when confronted with potentially adverse situations presumably lies exactly in the extent and quality of their stress responses to these situations (Kalisch et al., 2015), that is, in whether and in what way the potentially adverse situation really is a stressor to them. Clearly, we cannot afford to neglect this critical type of information. The only way out of this conundrum is that a resilience study must define a set of stimuli/situations that *potentially* are stressors for its subjects and must then assess whether a defined stimulus/situation has occurred or not.

An example of such an obvious design choice is a frequent type of trauma-resilience studies in which stressor measurement consists in registering a specific pre-defined type of potentially traumatizing life events (such as a physical accident or an act of violence) and which then assess the individual differences in mental health responses that occur in the aftermath of the event. Usually, in these studies, the sample only includes subjects in which the event has occurred in the first place. Studies that restrict stressor measurement to one strong potential stressor have the advantage that one can assume that the potential stressor really is a stressor to most study subjects. This means the study will most likely observe a substantial base rate of stressor-induced mental health problems, relative to which resilient responses can be defined, and researchers can be sure to address a significant mental health topic.

The downside of such a design choice is that one cannot exclude that the developing mental health problems that will be observed in some of the traumatized individuals originate from stressors other than the trauma, or at least only partly relate to the trauma. Resilience, then, in a traumatized subject may well be the consequence of the absence of such additional stressors, which were simply not measured in the study. This would obviously be a trivial explanation that does not necessitate any kind of protective, or resilience, processes. An example of a trauma resilience study where this is a probable explanation is a recent report that individuals' wealth predicts the absence of mental health problems following onset of physical disability (McGiffin et al., 2019). Wealth prevents or reduces exposure to additional stressors, such as the financial consequences of disability-related medical bills or a diminished household income and, thus, can be used to ease many of the pains and discomforts associated with the disability. Therefore, “resilient” subjects in the study may simply have been less exposed to other stressor than the disability itself.

We have argued earlier that treating a traumatic life event as an isolated stressor is to some extent artificial, as any major stressor is likely to be followed by other stressors (Kalisch et al., 2015; see also Norris and Elrod, 2006). For example, physical disability may cause financial problems, or the death of a close friend may compromise one’s social support network. We have therefore proposed that the measurement of stressor exposure in resilience studies should be extended to a wider range of potential stressors and we also argued that these should not only include potential other macrostressors but also seemingly minor events, such as daily hassles (Hahn and Smith, 1999). The necessity to also consider such microstressors is supported by long-standing evidence that these can also have profound negative mental health consequences, especially when they are many and long-lasting (Serido et al., 2004).

Extending stressor measurement to other macro- as well as to microstressors not only means an extension of the range of stressors to be measured but necessarily also implies a temporal extension. The stressful sequelae of a trauma or also the many different and temporally highly variable stressors associated with more chronic types of adversity or a stressful life transition phase cannot be assessed with a single measurement or over a very short time window only. In fact, the boundaries between traumatic event-like versus chronic adversity blur under this perspective, as is best illustrated by above example of the potentially traumatizing onset of a physical disability that mark a new life phase with significant chronic burden. Accurate stressor measurement thus requires repeated assessments over a reasonable time interval.

In MARP and LORA, we have chosen to assess macrostressors (LEs) and microstressors (DHs) retrospectively every three months over several years (Fig. 1). These monitoring intervals are arbitrary and constitute a trade-off between the needs to maximize sampling frequency and to minimize subject burden and costs. In other future studies, this trade-off may well be different. In order to limit burden and costs, the monitoring surveys are performed online in MARP and LORA. LE and DH item lists are exhaustive (27 and 58 items, respectively), to ideally not miss potential stressors, but they have also been pruned during their development to omit very infrequent and atypical stressors in modern Western societies (Chmitorz et al., in press; Canli et al., 2006). In contrast to the retrospective LE reports, subjects’ DH reports only cover the past week, because the accuracy of reporting minor stressors will decrease with the length of the reporting time window. We rely exclusively on self-report because we are not aware of a practicable non-subjective way of measuring stressors in a time-efficient manner. It is important to note that this makes our studies in principle vulnerable to confounding by reporting bias. For instance, subjects may be less inclined to report potential stressors that have objectively occurred to them but did not stress them noticeably; some subjects may have generally more accurate memories or a better verbal understanding of the lists than others; and some subjects may also be inclined not to report stressors that threaten their self-esteem. We hope to reduce such biases by the relatively short reporting time windows, by our instructions (which emphasize the necessity for objective and full reporting), and by the training in using the instruments that is effectively provided by the T0 report (which itself is not included in the analyses). Finally, we hope that the distinction we make in the instrument between the occurrence of an event or situation (first question for each item) and its subjective severity (second question to be answered only in case the first question was answered in the positive) encourages subjects to report the occurrence also of non-stressful events or situations.

The scoring of the results is explained in section iii.a) below.

**ii) How is mental health conceptualized and measured?**

We consider mental health to consist in the absence of symptoms, or mental dysfunctions. Rather than following the classical categorical system of mental disorders, whose suitability for mechanistic research has been questioned, we endorse a transdiagnostic approach that focuses on disturbances in functional systems of the brain/mind and investigates them dimensionally (Craddock and Owen, 2010; Cuthbert and Insel, 2013; see Kalisch et al., 2015). The four symptom clusters that constitute the four sub-scores of the GHQ-28 employed in MARP and LORA cover functional aberrations in the broad domains of aversive behavior, or defense (hypervigilance, worrying, and anxiety); social interaction (social withdrawal and other dysfunctions); appetitive behavior (despair, hopelessness, and depressed mood); and physiological regulation and bodily homeostasis (somatic tension, exhaustion, pain, and irritability), making the GHQ-28 suitable for transdiagnostic analysis. (Note, however, that in the current paper we focus on the analysis of the GHQ sum score.) As in the case of the LE and DH instruments, we emphasize the self-report nature of our mental health instrument. We currently do not have objective measurement systems for mental dysfunctions based on neurobiology and observable behavior that are both practicable in large-scale surveys and at the same time do not neglect the inherent subjective component of psychiatric symptoms. Like for LEs and DHs, the three-months monitoring interval is arbitrary and a study-specific trade-off between sampling frequency, subject burden, and costs.

**iii) What, concretely, constitutes a resilient outcome?**

Our outcome-based definition of resilience as maintenance or quick recovery of mental health during and after adversity (see above) implies that adversity is necessarily part of the equation. Only registering mental health outcomes without taking into account the adversity a subject was or is exposed to may be informative about mental health, but is not informative about resilience, which in its essence is mental health *despite* adversity (Mancini and Bonanno, 2009; Kalisch et al., 2017). This becomes clear from two hypothetical example scenarios. In example 1, two subjects in the MARP study both develop comparable moderate depressive symptoms over the course of their first two years of study at the local university. One of them, subject A, who moved to the city of Mainz from a far-away smaller city to study and since then struggled with the high costs of living in the region, did not manage to find friends among her peers, was left by her boyfriend (who did not tolerate the long-distance relationship), and failed important exams already in her first term. The other, subject B, was also left by her boyfriend, but already lived in the city before starting her studies and otherwise experienced no major difficulties. At the level of mental health, both subjects react similarly to the initial transition phase into higher education. But given that subject A was exposed to more stressors than subject B, it appears appropriate to classify A as more resilient than B.

In scenario 2, subject A develops much stronger depressive symptoms than subject B, approximately commensurate with her higher stressor exposure. Hence, at the level of mental health, subject A clearly reacts less favorably to her transition than subject B. However, taking into account the differences in stressor exposure now leads to the conclusion that both subjects exhibit similar resilience.

The scenarios raise the immediate question of when during an observation period a resilience outcome is to be registered. It may well be that someone shows initial deterioration of mental health under the pressure of acute challenges, but later, as pressure subsides, recovers. So, both subject A and B in our examples might leave the study after, for example, seven years in perfect mental health and it might then be concluded that they were both resilient to their life transition (although A perhaps more than B because A initially had more stressor exposure than B). However, if one had looked after two years, the picture would have been different.

To further complicate matters, it is unlikely that stressor exposure will be restricted to the beginning years of the life transition period that our subjects undergo. One can easily conceive yet another scenario 3, where subject A once more comes under severe pressure towards the end of her university studies, perhaps related to fear of a big final exam, or due to poor professional perspectives in her field of study, or to the onset of a physical disease. This would make her difficult to compare to subject B, who, in this scenario, would again be largely spared from any major adversity. One would also wonder if any deterioration in mental health that A shows under the immediate influence of renewed adversity is lasting or only temporary (such that A will recover in the period after the conclusion of the study).

More generally, the challenge in studying resilience to most types of stressors lies not only in the need to measure potential stressors exhaustively and at high frequency and to accurately quantify them (discussed in i), but also in the complex temporal patterns in which they can occur. Life is stressful. This complication appears even more in the LORA study, in which no specific life transition phase is addressed. But it is a general problem faced by any longitudinal study that attempts to realistically assess stressor exposure and associated mental health changes in the way discussed above. The complication is also related to the sheer variety of stressors humans can experience, meaning that one may experience one type of stressor during one period of one's live and another during another period, and this in a multitude of possible combinations of stressors and exposure periods. This makes it difficult (or even impossible) to apply a simple formula for quantifying a person's resilience. For instance, one might be tempted to assess mental health during the beginning and the end of a study observation period (in MARP, this would be, for instance, years 1 and 7), sum up total stress exposure during the whole observation period, and then simply relate this to changes in mental health (values at year 7 minus values at year 1).

Clearly, this would be problematic. Taking scenario 3, subject A's mental health report at year 7 would presumably be strongly affected by her recently experienced adversities and only to some extent reflect the integral of stressor exposure, mental health, and subjective quality of life over the entire seven years of her study participation. As a result, A would score with strong mental health deterioration from year 1 to year 7. One can easily conceive a third subject C with the quantitatively same stressor exposure as A, but at a different – earlier – time during the study, such that C would already have recovered in year 7. C would then score with less year 1-to-year 7 mental health deterioration than A and be classified as more resilient than A, given that both show the same summed stressor exposure.

Given these caveats and pitfalls, how can we apply the formula proposed by Kalisch et al. (2015) for the quantification of resilience (Fig. 2) – resilience being understood in conformity with the consensus definition in Kalisch et al. (2017) as a good mental health outcome following stressor exposure – to studies performing longitudinal stressor exposure and mental health monitoring? Above considerations call for an approach that takes into account the temporal patterns of stressor exposure and mental health changes in more detail and only then aggregates the information into summary quantities. Our proposal for such an approach is described in the following.

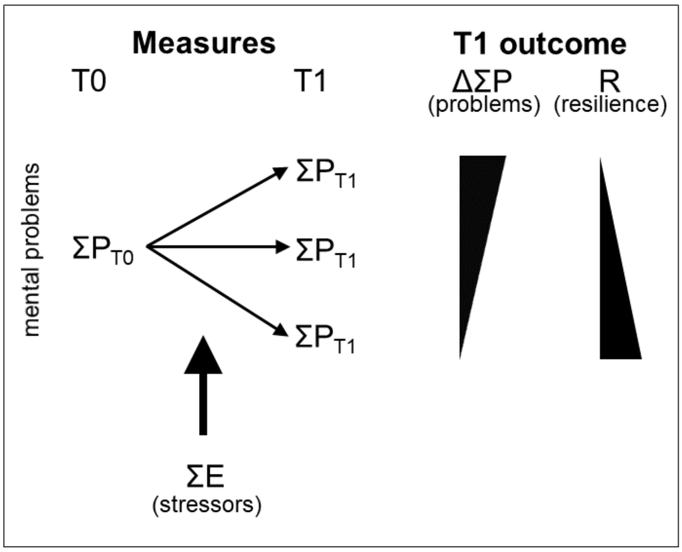


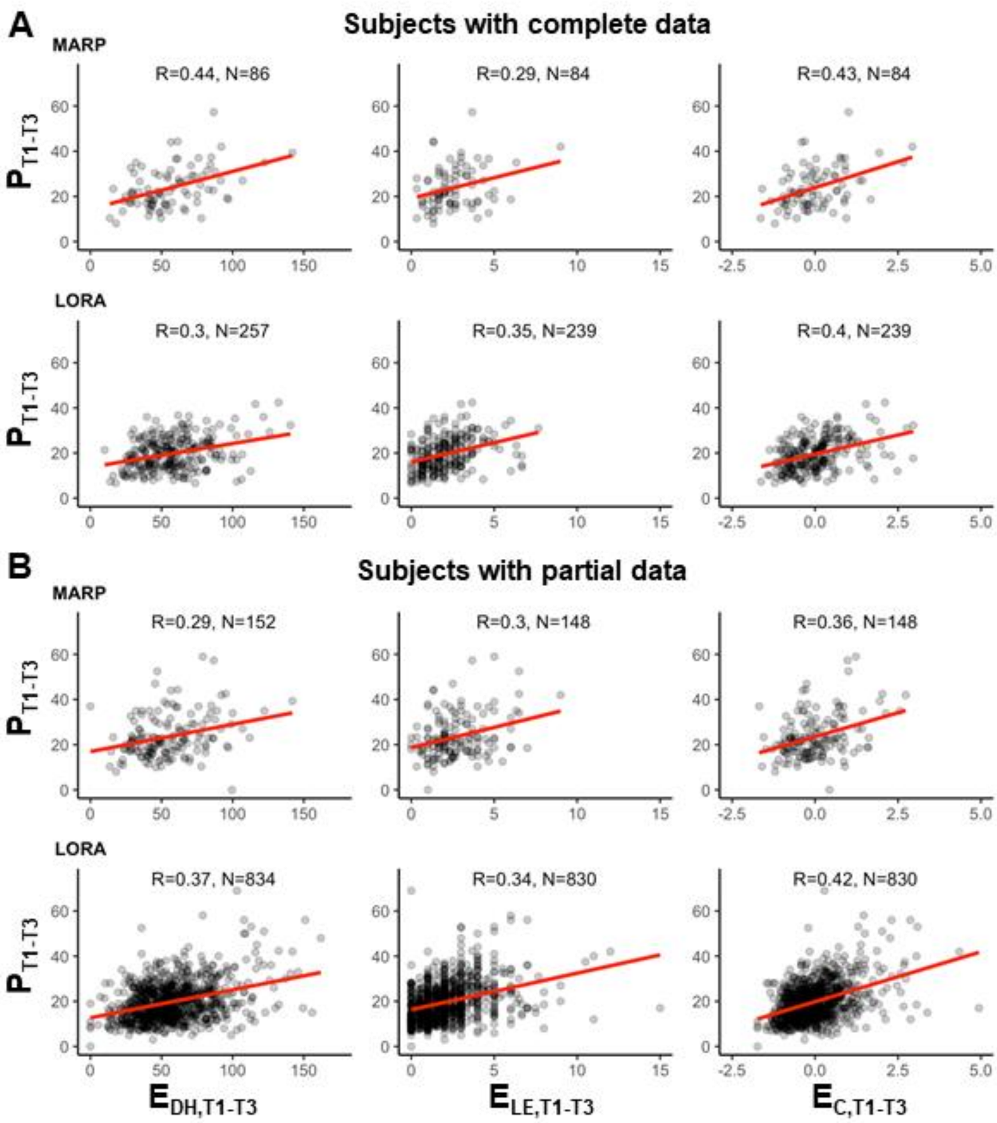

**Figure 2. Outcome-based quantification of resilience.**

Kalisch et al. (2015) proposed to quantify resilience (R) as the ratio of changes in mental health problems P from before (T0) to after (T1) stressor exposure E. Stressors are summed over the observation period (from T0 to T1). Adapted from Kalisch et al., 2015 (their Fig. 1). With permission from Cambridge University Press.

### iii.a) Residualization-based calculation of stressor reactivity

Figure 3 shows the relationship between the self-reported exposure to LEs ($E_{LE}$) as well as DHs ($E_{DH}$) and subjects' self-reported mental health problems (P) in both MARP and LORA during the first nine months of the study. The maximum possible LE occurrence count obtained at any three-monthly monitoring time point is 27, according to the number of list items. The maximum three-monthly DH count is the number of list items (58) times the maximum number of past days on which a DH can have occurred (7), i.e., 406. We average both LE and DH counts over three consecutive monitoring time points, starting with the first monitoring time point (T1) after study inclusion (T0) (cf. Fig. 1), to thus obtain average T1-T3 counts ($E_{LE,T1\text{-}T3}$; $E_{DH,T1\text{-}T3}$). Averaging is beneficial in the case of DHs because any monitoring time point only covers DH occurrence in the past week, a time window that might be burdened with too much variability for being a useful representative for the entire preceding three-months period. Averaging is also advantageous in the case of LEs, given their rare occurrence. Accordingly, we also average GHQ scores to obtain $P_{T1\text{-}T3}$. We here only report the total GHQ-28 sum scores but not the sub-scores, both for simplicity and because we are primarily interested in discovering general resilience mechanisms, i.e., mechanisms that protect not only against single, but several, mental dysfunctions (Kalisch et al., 2015).

In both samples, we find the E-P relationships to be predominantly monotone (with no strong deviations from linearity) and in the expected range; Fig. 3). This applies when only including subjects that have provided all three reports for both variables (Fig. 3A; subjects with complete data) and when also including subjects that have only provided a minimum of one report in each variable (Fig. 3B; subjects with partial data). The latter allows us to focus all further analyses on linear relationships in subjects with partial data only, thereby increasing power. We also calculate a combined LE/DH stressor exposure score $E_{C,T1\text{-}T3}$ as the mean of the z-scores of the DH and LE counts. This also seems appropriate as LE and DH counts are




well correlated in both samples (MARP: R=0.38, p=2e-06, N=148; LORA, subjects with partial data: R=0.38, p<2.2e-16, N=702) and because of the theoretically close relationship between macro- and microstressors for which we argued above (section i). The $E_C$-P relationships are consistently more stable than either $E_{LE}$-P or $E_{DH}$-P relationships, which is why we also focus all further analyses on the combined LE/DH exposure score $E_C$.

**Figure 3. Relationship between stressor exposure (E) and mental health problems (P) in MARP and LORA.**

LE and DH counts as well as GHQ total sum scores are averaged over the first 3 three-monthly monitoring time points after study inclusion (T0), that is: T1 (month 3), T2 (month 6), and T3 (month 9), to obtain stressor exposure scores $E_{DH,T1\text{-}T3}$ (left column) and $E_{LE,T1\text{-}T3}$ (middle column) and mental health problem score $P_{T1\text{-}T3}$. $E_{C,T1\text{-}T3}$ is a combined stressor exposure score (mean of $E_{LE}$ and $E_{DH}$ z-scores; right column). Correlations in (A) are from subjects with complete data (all three monitoring time points) for all variables in a given analysis. Correlations in (B) are from subjects with at least one time point per variable.

Given the observation in two independent samples of a robust monotone $E_C$-P relationship, we here propose that the distance of an individual's P score to the regression line is likely to be informative about the reactivity of his/her mental health to stressor exposure in the covered time interval. Figure 4 illustrates the principle of residualizing individual mental health problem scores P on the regression line defined by the group's $E_C$-P relationship. The regression line is the normative reactivity of mental health to stressor exposure (in short: 'stressor reactivity', SR) in the whole group during the T1-T3 time window. A subject's residual expresses to what extent he or she deviates from that normal $E_C$-P relationship. Individuals with positive residual values (red dots above the line in Fig. 4) show "too many" mental health problems, given their level of stressor exposure; individuals with negative values (green dots below the line in Fig. 4) show "too few" mental health problems, given their stressor exposure (ignoring random variability for the moment). In other words: a positive residual reflects over-reactivity of mental health to stressor exposure (high stressor reactivity, high SR) compared to the sample average; a negative residual reflects under-reactivity reactivity (low SR). Residualization has been introduced to the resilience literature by Amstadter et al. (2014) and van Harmelen et al. (2017).

The residualization approach has the advantage that it inherently corrects for individual differences in the level of stressor exposure. In Figure 4, the same mental health problem score P in two arbitrarily chosen subjects 1 and 2 does not constitute equal degrees of stressor reactivity SR, as subject 1 has experienced fewer stressors than subject 2. That is, subject 1 is actually more reactive to stressors than subject 2 despite identical P scores. Subject 1's stressor reactivity, however, is comparable to that of subject 3, who has a higher mental health problem score P but also a higher level of stressor exposure E. Accordingly, subject 1's and 3's residuals (thick red lines) are comparable. Clearly, all subjects lying above the regression line are more stressor reactive than those subjects below the line. Subjects 4 and 5 in the figure have comparable below-average stressor reactivity.

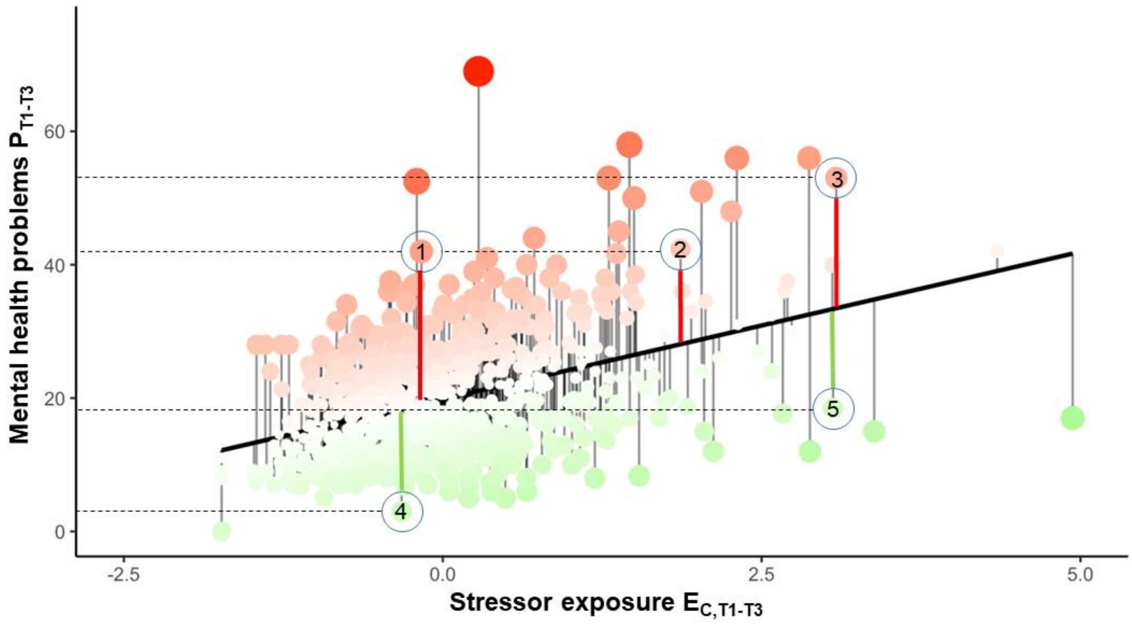



**Figure 4. Individual mental health reactivity to stressor exposure ('stressor reactivity').**

The regression line shows the normative linear positive relationship between combined exposure to LEs and DHs stressors ($E_{C,T1\text{-}T3}$) and mental health problems ($P_{T1\text{-}T3}$) in N=830 LORA subjects providing partial data during the first nine months of the study (monitoring time points T1-T3). The residuals onto the regression line are subjects' deviations from the normative $E_C$-P relationship. A strong positive deviation reflects high susceptibility of the subject's mental health to the effects of DHs and LEs (high stressor reactivity, SR) during the chosen time window; a strong negative deviation reflects below-average low susceptibility (low SR). 1 to 5 denote arbitrarily chosen subjects.

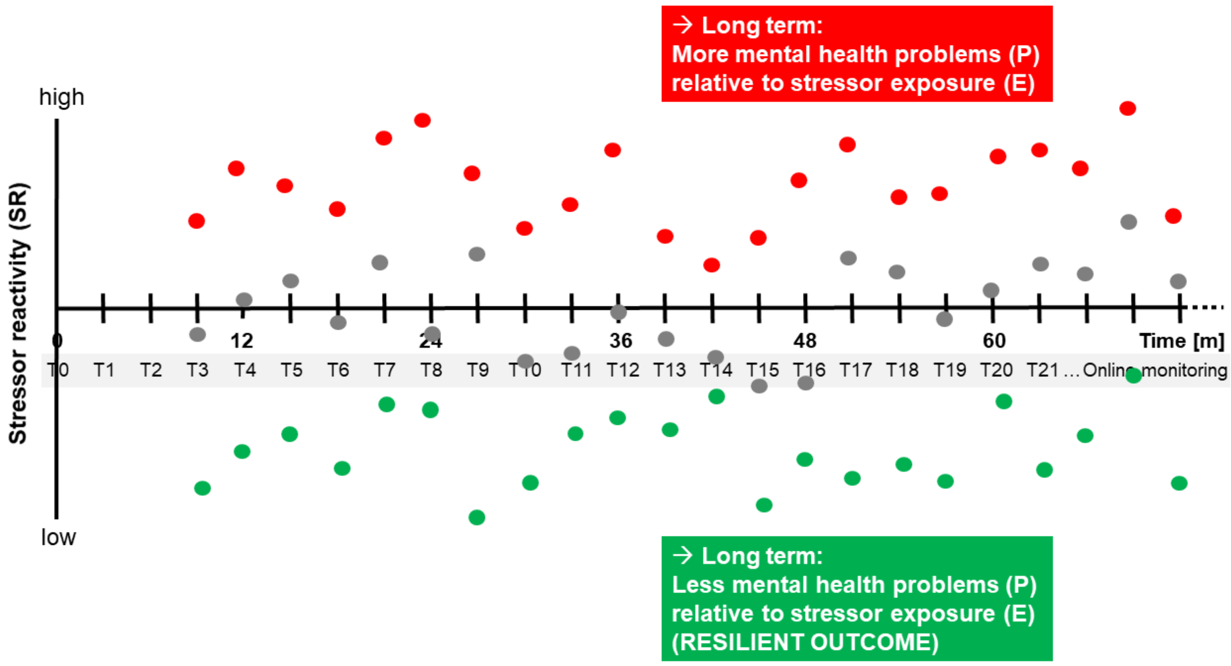
### iii.b) Time courses of stressor reactivity

Low stressor reactivity during T1-T3 can be interpreted as a first indication of a subject's resilience, in line with our definition of resilience as an outcome, that is, as long-term mental health maintenance despite stressor exposure. A subject with low stressor reactivity has accumulated less mental health problems across the first nine study months than a subject with high stressor reactivity, while differences in stressor exposure are controlled for. Of course, nine months is an arbitrarily chosen time span and hardly reflects a long-term outcome, though by increasing the length of the time window, one could make the corresponding SR score more closely approximate a true long-term outcome.

An individual with consistently low stressor reactivity over a longer time span will be mentally healthier despite stressors over that time span than an individual with consistently high stressor reactivity, provided comparable stressor exposure.

Hence, one possibility would be to build SR scores based on time windows considerably longer than T1-T3. This, however, would neglect a potential - or even likely - dynamic aspect of stressor reactivity. While it is reasonable to assume that an individual's stressor reactivity is influenced by trait-like individual characteristics and therefore exhibits some stability over time, it is also safe to assume that one's stressor reactivity may to some extent be affected by non-stationary internal or environmental factors and that may not be constant over time. This is plausible given increasing evidence that traits, too, can change over time (Soto et al., 2011) and are behaviorally expressed in a contextually dependent and variable way (Matthews, 2018). Thus, stressor reactivity is likely to show meaningful temporal variation – which would be overlooked by averaging over more measurement time points. Another argument that speaks in favor of a more dynamic approach is the temporally variable nature of stressor exposure discussed above. Therefore, for long-term studies such as MARP and LORA, we propose to define individual SR time courses rather than integrating SR over an entire study time window.

We do so in sliding windows of temporally overlapping SR scores (T1-T3; T2-T4; T3-T5; …). This approach allows for describing potential temporal fluctuations in stressor reactivity. At the same time, the use of (moderately) extended and overlapping windows effectively introduces temporal smoothing of the SR time courses and thus reduces spurious fluctuations. The schematic graph in Figure 5 shows SR time courses in three hypothetical LORA subjects, one with consistently high stressor reactivity (red), one with consistently normal reactivity (grey), and one with consistently low reactivity (green). After several years of monitoring in the study, the "red" subject will have accumulated more mental health problems relative to his or her stressor exposure than the "grey" and "green" subjects. The grey and green subjects




then would be considered more resilient than the red subject, in accordance with the formula for resilience quantification proposed by Kalisch et al. (2015; see also Figure 2) and with the general outcome-based definition of resilience as consensually introduced by Kalisch et al. (2017).

**Figure 5. Hypothetical time courses of stressor reactivity.**

Three hypothetical study subjects with consistently high (red), average (grey) and low (green) stressor reactivity (SR), determined in sliding windows of three consecutive monitoring time points (T1-T3; T2-T4; T3-T5; …), are shown. In the long run, consistently lower-than-normal stressor reactivity leads to fewer mental health problems P relative to individual stressor exposure E, that is, to a more resilient outcome.

Importantly, however, with SR time courses there is no single time point in the study at which a final resilience outcome is calculated. We thereby respond to the problem highlighted above (iii) that stressor exposure in real life is usually not temporally circumscribed and may occur at different times and in different forms in different study subjects. That is, it is impossible in most studies to define a clear time window of stressor exposure relative to which changes in mental health and, thus, a resilience outcome could be determined. This can be considered another advantage that building SR time courses has over a non-dynamic approach.

Based on the SR time course, still aggregated quantities might be calculated, but the original temporal pattern should always be kept as a backdrop. While also alternatives to the proposed sliding-window approach could be devised (e.g., based on differential equations that directly model the impact of stressors on mental health), the dynamic nature of SR should always be considered.

**iv) What are the potential adaptation processes and how are they measured?**

If there is no single way of determining a final resilience outcome in all subjects, this raises the question of what dependent variable can be used to identify baseline (T0) predictors of resilience (in the sense of 'resilience factors' making a resilient outcome more likely) or processes of adaptation ('resilience processes' that effectively generate the outcome). For baseline resilience factors, we define the first SR score ($SR_{T1\text{-}T3}$ ) as the outcome variable for calculating a preliminary prediction analysis in MARP and LORA (Fig. 6A). We are aware that this is a short prediction time window, and we will therefore include additional SR scores (T2-T4; T3-T5; …) into the analysis as the studies progress, e.g., using linear mixed effects models to take into account repeated measurements.

Our emphasis is, however, on the identification of resilience processes. We here take advantage of the on-site testing batteries that are regularly repeated in the MARP and LORA subjects (see Fig. 1). These batteries provide time-dependent predictor variables (X in Figure 6B), i.e., time courses of predictor variables that can potentially explain considerably more variance in SR time courses than a single measurement of any predictor variable at baseline. Time-dependent battery variables that explain significant variance in the SR time courses can then be seen as pointing towards allostatic adaptations in those functions or properties of the brain/mind, the body or the environment that they index.

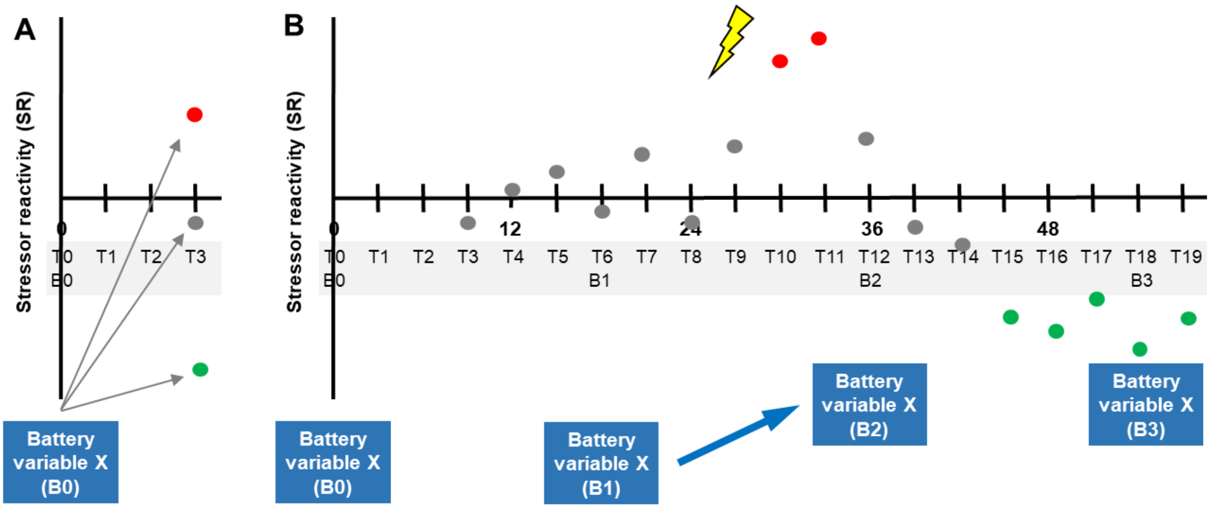



**Figure 6. Potential relationships between testing battery measurements and time courses of stress-or reactivity.**

(A) Prediction of future SR (here $SR_{T1\text{-}T3}$) by baseline (B0) battery variable X. (B) Hypothetical scenario where, under the influence of significant stressor exposure (lightning bolt), an individual with normal stressor reactivity (grey) improves on a battery measurement X of a resilience factor from B1 to B2 (indicated by the arrow) and where his/her SR decreases (turns green) after an initial increase evoked by the acute stressor effect. This illustrates a lasting change in how the system copes with adversity.

To illustrate the distinction between resilience factors and processes and their relation to SR time courses, we reiterate a simple scenario employed in Kalisch et al. (2019) for that purpose. We assume that volitionally regulating emotions away from negative towards more positive emotional states using verbal strategies of reappraising the meaning of, or reframing, potentially threatening situations ("positive reappraisal"; Lazarus and Folkman, 1984; Gross, 1998) protects mental health in stressor-exposed individuals. Reappraisal exerts this assumed effect by dampening stress reactions and thereby limiting the expense of resources (time, energy, cognitive capacity, financial or social capital) that accompanies them and that can be deleterious to body and mind if the stress reactions are too intense and too frequent. Hence, every time an individual is confronted with a significant stressor that more or less automatically induces a negative emotional state and then reappraises the situation in a benign way (while also avoiding unrealistically positive appraisals), he or she will save resources and make mental health impairments less likely to occur (Kalisch et al., 2015). (In the terminology introduced in Kalisch et al. (2015), the concrete act of performing reappraisal is a resilience mechanism.)

Individuals with a good ability and habitual tendency to use reappraisal for emotion regulation are more likely to regulate their emotions using reappraisal (to employ the resilience mechanism that they have at their disposal), and it can be hypothesized that any good measure of reappraisal usage taken at T0 will therefore negatively predict $SR_{T1\text{-}T3}$ and later SR time windows. (Such measures may be a questionnaire or a laboratory reappraisal task or perhaps an index of neural activity associated with a reappraisal task.) That is, better reappraisers will be less stressor reactive during the first nine or more study months. If this is correct, individual reappraisal ability/tendency can be considered a resilience factor.

If the measurement of reappraisal usage (resilience factor) is repeated regularly (such as with every online monitoring or, as is practiced in MARP and LORA, with every application of the testing battery B0, B1, B2, …), this can also inform about reappraisal as a process. Two resulting patterns are conceivable. An individual may not change his or her reappraisal ability/tendency much over time, that is, may just stay a more or less good or bad reappraiser. Good reappraisers should show less stressor reactivity than bad reappraisers also at monitoring time points past T3, and the inter-individual variance of SR time courses explained by the repeatedly measured reappraisal ability/tendency should increase, relative to a single measurement of reappraisal ability/tendency only at T0, simply because repeated measurements provide more reliable information on individual reappraisal ability/tendency. From such an apparent dampening effect of reappraisal usage on stressor reactivity in good reappraisers one can then infer, indirectly, that these individuals presumably successfully use reappraisal in their daily lives (in the sense of a process) and that this suffices them to maintain emotional balance and good mental health. That is, the external perturbations that the system experiences are not strong enough to exceed the system's capacity for coping, and, consequentially, the system does not lastingly change its way of functioning. The individual copes as it is used to cope. In the terminology introduced at the start of this paper, coping in this case is homeostatic, and real-life use of reappraisal is a homeostatic resilience process.

The situation is different when an individual's reappraisal capacity is not sufficient to maintain stability. In this case, the system will have to adapt in an allostatic way, i.e., it will struggle to find new or better ways of coping that constitute lasting shifts in system function. For example, important life challenges may be answered by greater reappraisal efforts and, if successful, this may make an individual a better reappraiser than he or she was while life was still less challenging (see the scenario in Kalisch et al., 2019, their Fig. 7). In the testing battery employed in MARP and LORA, such improvements should become apparent in higher values on the reappraisal ability/tendency predictor variable X, along with the variable well explaining variance in SR. In Figure 6B, a hypothetical time course is shown where, under the influence of significant stressor exposure (lightning bolt), an individual with normal stressor reactivity (grey) improves on reappraisal ability/tendency (battery variable X) and where his or her stressor reactivity decreases (turns green) after an initial increase evoked by the acute stressor effect. Thus, the example illustrates a lasting change in a measure of a resilience factor that indicates an allostatic process of adaptation. Hence, as opposed to the above example of real-life reappraisal use as a homeostatic resilience process, allostatic resilience processes occur at a higher level of system function, by constituting shifts in the way the system operates.

It is, of course, also conceivable that an important life challenge will not be answered by the organism with an increase specifically in reappraisal ability/tendency, but maybe in some other coping function or mechanism. This would also be an allostatic resilience process and should become apparent in the testing battery, provided the existence of battery variables that index that function.

Finally, the system may also not succeed in adapting well to the situation. This should become apparent in lasting increases in stressor reactivity as well as potentially in decreases in testing battery variables that index resilience factors. (One might give up using reappraisal and become an even worse reappraiser than before; in turn, variables associated with pathological forms of coping, e.g., catastrophizing or rumination, might increase.) The latter scenario would also constitute an allostatic process, but a maladaptive one, which should accordingly be classified as pathological or pathogenic, as opposed to the adaptive resilience processes discussed before.

Hence, a combination of repeated measurements of stressor reactivity with repeated battery measurements of system functions (including hypothesized resilience factors) can help identify both homeostatic and allostatic resilience processes.

**v) How can the statistical relationship between adaptation processes and resilience as an outcome be established?**

Next to the use of reliable measures, the successful identification of resilience processes will depend on the availability of suitable mathematical methods to link resilience predictors and outcomes. There are multiple ways to assess the statistical relationship between changing individual properties, as measured in battery variables X, and resilience, as assessed using changing SR scores. A rather simple approach would draw on the literature on dynamic predictions in clinical settings (Putter, 2013) and apply it to the smoothed SR scores. Such an approach might predict the following SR scores either with the latest baseline information or a function of all previous baseline information, which allows for taking the history of adaptation processes into account. More specifically, each smoothed mean SR score (T1-T3; T2-T4; T3-T5; …) is predicted with (the latest) baseline information of B0, B1, or later. The temporal sequence of regression coefficients $\beta_{1,\text{T1-T3}} - \beta_{1,\text{T3-T5}}$ of the same predictor X can be connected and, for instance, be smoothed to reflect the idea of slow changes. Statistical significance can be assessed with appropriate standard errors taking the repeated measures into account. Algorithms for variable selection can assist the search for resilience factors in big baseline batteries (Schmidtmann et al., 2014; Zöller et al., 2016).

In case the resilience factor changes over time, its prediction capabilities might be reduced the more time lies between the monitoring time point T and the latest measurement battery B. This "aging" of predictors can be attenuated by modeling the trajectories of the resilience factors, potentially taking the stressor load and mental health into account.

Regression-based approaches are limited in that they assume uni-directional causality (resilience factors influence the SR score but not vice versa). However, there are other models which can treat each single observation of P, E, and the battery assessments of potential resilience factors X as samples from continuous trajectories. Such dynamic models, which can, for instance, be based on differential equations, take into account the (inferred) value of P, E, and the resilience factors at every point in time. They require and benefit from domain expertise, since every trajectory affects the change of itself and all other variables according to a predefined system of equations. Their estimated coefficients can also be tested for significance (Raue et al., 2009). Such models allow for irregularly sampled measures. Accordingly, they are able to bridge more disparate temporal resolutions as well as entirely missing observations. Yet, methods development for such approaches is still on-going, i.e., they are not yet readily available.

## 3. Limitations and comparison with other approaches

A potential criticism of stressor reactivity scores might be that they are also calculated for individuals with very low stressor exposure E (left end of the regression line in Figure 4), and that resilience in the absence of significant adversity is not a meaningful concept (Mancini and Bonanno, 2009). This criticism, however, would fail to consider that stressor exposure may well change over the course of a study and that a subject with initially low stressor exposure might well experience more stressors later. This means that excluding a subject from the calculation of stressor reactivity scores based on a low E score at a given time point would unnecessarily exclude that subject from the long-term observation needed to determine resilience. Further, given the monotone, practically linear E-P relationships in our studies (Fig. 3), any such decision to exclude a subject would have to be based on an E threshold that is not anchored to any kind of turning point or “true” threshold present in the stressor exposure data.

One could still argue that subjects showing consistently low E scores over the course of a study may better be excluded, to thus fulfill the reasonable criterion that adversity has to be present. We have therefore taken the approach in our analyses of MARP and LORA data that a) primary analyses based on the entire samples be complemented by secondary analyses of those two thirds of the subjects with the highest overall stressor exposure E, and that b) the results of those secondary analyses should go in the same direction as those of the primary analyses, in order for the primary analysis results to be considered valid.

Another criticism may be the combination of micro- and macrostressors in a common $E_C$ score (iii.a). Although the momentary impact of any given microstressor is presumably different from the impact of any macrostressor, micro- and macrostressors are tightly related, as described above. Yet there is, to this date, no universal quantitative way to reliably describe their relation, let alone to do so in a longitudinal fashion. Hence, using the mean of the z-scores of the DH and LE counts may as of now be the best practical solution, which is empirically supported by our finding that $E_C$-P relationships are consistently more stable than either $E_{LE}$-P or $E_{DH}$-P relationships.

To exclude potential misunderstandings, we would like to emphasize that stressor reactivity-based trajectories differ fundamentally from the type of mental health trajectories that have been prominently used in resilience research to delineate mental health responses to potentially traumatizing events (e.g., loss of a spouse, stroke, …) or to onset of chronic adversity (e.g., physical disability, chronic pain, …) (see above and Bonanno et al., 2011). Although in some cases these studies have succeeded in controlling for the level of stressor exposure at the start of the trajectory (by assuring that the severity of the traumatic event or of the chronic adversity whose onset defines trajectory start is comparable across the cohort), they nevertheless do not take into account stressors occurring after trajectory start. As discussed above, such stressors may include the potentially individually very different sequelae of a trauma or some chronic type of adversity resulting from the trauma. These studies can therefore not exclude that shifts in mental health often observed in these cohorts, whether from good to bad or vice versa, are caused merely by individual differences in stressor exposure. By contrast, stressor reactivity trajectories as introduced here inherently take such potential influences into account.

In a recent paper, we have proposed a general framework for an approach to determine resilience processes based on a dynamic network account of psychiatric disorders and resilience (Kalisch et al., 2019). Both that and the current account converge in that they rely on frequent repeated measurements of stressor exposure, mental health problems and potential resilience factors, and in that they consider these to be usually time-variant. The two approaches can thus be seen as complementary. The network account is, however, considerably more conceptual than the current account and awaits concrete mathematical formulation. In particular, non-trivial problems related to robust parameter estimation and reliable determination of time-dependent contemporaneous and time-lagged correlations between the network nodes representing psychiatric symptoms and resilience factors have to be solved, especially where symptoms and resilience factors are measured at discrete time points or on different time scales. Also, other approaches for dynamic modeling, such as differential equations, might be considered, yet have still to be fully worked out. By contrast, the proposed sliding-window residual approach affords a way to already now test resilience factors and resilience processes in existing data.

## 4. Outlook

As resilience research is moving from studying resilience-conducive traits to studying malleable and time-dependent resilience factors and from identifying baseline resilience predictors to characterizing processes of adaptation, the development of suitable study designs, analytical concepts, and associated mathematical methods becomes a crucial field of methodological research. The present paper aims to enrich the current debate and to propose a concrete solution, which we believe proposes important novel elements: a study scheme involving the high-frequent concurrent measurement of micro- and macrostressors in combination with repeated measurements of mental health and potential resilience factors; a way to quantify resilience as a dynamically changing outcome; and a way to link resilience factors with the outcome in a dynamic fashion, to thus identify resilience processes. We consider our proposal a generic solution that can serve as a blueprint for future resilience studies, notwithstanding necessary adaptations to the concrete context (e.g., by choosing other measurement and sliding-window intervals, by using stressor assessment instruments that are better suited for the population of interest, or by measuring mental health with different tools). We are confident that this and other dynamic approaches will considerably advance the field of resilience research.

## References


Amstadter AB, Myers JM, Kendler KS (2014) Psychiatric resilience: longitudinal twin study. Br J Psychiatry 205:275–280.

Bonanno GA (2012) Uses and abuses of the resilience construct: Loss, trauma, and health-related adversities. Soc Sci Med 74:753–756.

Bonanno GA, Romero SA, Klein SI (2015) The Temporal Elements of Psychological Resilience: An Integrative Framework for the Study of Individuals, Families, and Communities. Psychol Inq 26:139–169.

Bonanno GA, Westphal M, Mancini AD (2011) Resilience to Loss and Potential Trauma. Annu Rev Clin Psychol 7:511–535.

Canli T, Qiu M, Omura K, Congdon E, Haas BW, Amin Z, Herrmann MJ, Constable RT, Lesch KP (2006) Neural correlates of epigenesis. Proc Natl Acad Sci USA 103:16033–16038.

Chmitorz A, Kurth K, Mey LK, Wenzel M, Lieb K, Tüscher O, Kubiak T, Kalisch R (in press) Mainz Inventory of Microstressors (MIMIS): Development and ecological validation of a microstressor scale for adults. JMIR Mental Health. doi: 10.2196/14566.

Compas BE, Wagner BM, Slavin LA, Vannatta K (1986) A prospective study of life events, social support, and psychological symptomatology during the transition from high school to college. Am J Commun Psychol 14:241–257.

Craddock N, Owen MJ (2010) The Kraepelinian dichotomy - going, going... but still not gone. Br J Psychiatry 196:92–95.

Craske MG, Wolitzky–Taylor KB, Mineka S, Zinbarg R, Waters AM, Vrshek–Schallhorn S, Epstein A, Naliboff B, Ornitz E (2012) Elevated responding to safe conditions as a specific risk factor for anxiety versus depressive disorders: Evidence from a longitudinal investigation. JAbnorm Psychol 121:315–324.

Cuthbert BN, Insel TR (2013) Toward the future of psychiatric diagnosis: the seven pillars of RDoC. BMC Med 11:126.

Eisenberg D, Hunt J, Speer N (2012) Help seeking for mental health on college campuses: review of evidence and next steps for research and practice. Harv Rev Psychiatry 20:222–232.

Goldberg DP, Hillier VF (1979) A scaled version of the General Health Questionnaire. Psychol Med 9:139–145.

Gross JJ (1998) Antecedent- and response-focused emotion regulation: Divergent consequences for experience, expression, and physiology. J Pers Soc Psychol 75:224–237.

Hahn SE, Smith CS (1999) Daily hassles and chronic stressors: conceptual and measurement issues. Stress Med 15:89–101.

Hammen C (2016) Depression and stressful environments: identifying gaps in conceptualization and measurement. Anx Stress Cop 29:335–351.

Herbst U, Voeth M, Eidhoff AT, Müller M, Stief S (2016) Studierendenstress in Deutschland- eine empirische Untersuchung. AOK-Bundesverband.

Kalisch R et al. (2017) The resilience framework as a strategy to combat stress-related disorders. Nat Hum Behav 1:784.

Kalisch R, Cramer AOJ, Binder H, Fritz J, Leertouwer Ij, Lunansky G, Meyer B, Timmer J, Veer IM, van Harmelen A-L (2019) Deconstructing and Reconstructing Resilience: A Dynamic Network Approach. Perspect Psychol Sci 14:765–777.

Kalisch R, Müller MB, Tüscher O (2015) A conceptual framework for the neurobiological study of resilience. Behav Brain Sci:e92.

Karatsoreos IN, McEwen BS (2011) Psychobiological allostasis: resistance, resilience and vulnerability. Trends Cogn Sci 15:576–584.

Kendler KS, Gardner CO (2010) Dependent stressful life events and prior depressive episodes in the prediction of major depression: the problem of causal inference in psychiatric epidemiology. Arch Gen Psychiatry 67:1120–1127.
Kent M, Davis MC, Reich JW (2014) Introduction. In: The resilience handbook, pp xii–xix. New York, USA: Routledge.
Lazarus RS, Folkman S (1984) Stress, appraisal and coping. New York, USA: Springer.
Luthar SS, Cicchetti D, Becker B (2000) The construct of resilience: a critical evaluation and guidelines for future work. Child Dev 71:543–562.
Mancini AD, Bonanno GA (2009) Predictors and Parameters of Resilience to Loss: Toward an Individual Differences Model. J Pers 77:1805–1832.
Matthews G (2018) Cognitive-Adaptive Trait Theory: A Shift in Perspective on Personality. J Pers 86:69–82.
McEwen BS SE (1993) Stress and the individual: Mechanisms leading to disease. Arch Intern Med 153:2093–2101.
McGiffin JN, Galatzer-Levy IR, Bonanno GA (2019) Socioeconomic resources predict trajectories of depression and resilience following disability. Rehabil Psychol 64:98–103.
Norris FH, Elrod CL (2006) Psychosocial consequences of disaster: a review of past research. In: Methods for disaster mental health research, pp 20–44. New York, USA: The Guildford Press.
Paykel ES (2003) Life events and affective disorders. Acta Psychiatr Scand Suppl:61–66.
Putter H (2013) Landmarking. In: Handbook of survival analysis. Boca Raton, USA: Chapman & Hall/CRC.
Raue A, Kreutz C, Maiwald T, Bachmann J, Schilling M, Klingmüller U, Timmer J (2009) Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. Bioinform 25:1923–1929.
Reavley N, Jorm AF (2010) Prevention and early intervention to improve mental health in higher education students: a review. Early Intervent Psychiatry 4:132–142.
Rotenstein LS, Ramos MA, Torre M, Segal JB, Peluso MJ, Guille C, Sen S, Mata DA (2016) Prevalence of Depression, Depressive Symptoms, and Suicidal Ideation Among Medical Students: A Systematic Review and Meta-Analysis. J Am Med Assoc 316:2214–2236.
Rutter M (2012) Resilience as a dynamic concept. Dev Psychopathol 24:335–344.
Sapienza JK, Masten AS (2011) Understanding and promoting resilience in children and youth. Curr Opin Psychiatry 24:267–273.
Schmidtmann I, Elsäßer A, Weinmann A, Binder H (2014) Coupled variable selection for regression modeling of complex treatment patterns in a clinical cancer registry. Stat Med 33:5358–5370.
Seery MD, Alison E, Silver RC (2010) Whatever does not kill us: Cumulative lifetime adversity, vulnerability, and resilience. J Pers Soc Psychol 99:1025–1041.
Serido J, Almeida DM, Wethington E (2004) Chronic stressors and daily hassles: unique and interactive relationships with psychological distress. J Health Soc Behav 45:17–33.
Soto CJ, John OP, Gosling SD, Potter J (2011) Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. J Pers Soc Psychol 100:330–348.
Sterling P, Eyer J (1988) Allostasis: a new paradigm to explain arousal pathways. In: Handbook of life stress, cognition and health, pp 629–649. New York: Wiley.
van Harmelen A-L, Kievit RA, Ioannidis K, Neufeld S, Jones PB, Bullmore E, Dolan R, NSPN Consortium, Fonagy P, Goodyer I (2017) Adolescent friendships predict later resilient functioning across psychosocial domains in a healthy community cohort. Psychol Med 47:2312–2322.
Zöller D, Schmidtmann I, Weinmann A, Gerds TA, Binder H (2016) Stagewise pseudo-value regression for time-varying effects on the cumulative incidence. Stat Med 35:1144–1158.

**Acknowledgements**
This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 777084 (DynaMORE project), from the Deutsche Forschungsgemeinschaft (DFG grant CRC 1193, subprojects B01, C01, C03, C04, C05, C06, C07, Z03), from the State of Rhineland-Palatinate (Stiftung Rheinland-Pfalz für Innovation, DRZ programme), and from the Leibniz Institute for Resilience Research.