Adversity and Resilience Science
https://doi.org/10.1007/s42844-022-00078-6

ORIGINAL ARTICLE

Quantifying Resilience as an Outcome: Advancing the Residual Approach with Influence Statistics to Derive More Adequate Thresholds of Resilience

J. Höltge[1] · M. Ungar[1]

Accepted: 14 September 2022

Abstract
Resilience as an outcome is defined as better-than-expected wellbeing and developmental progress in the context of exposure to significant adversity. This definition has been quantified through a process of residualization, the difference between an individual's observed outcome score and its expected (predicted) outcome score using statistical modeling. This approach can be biased by the presence of individuals that disproportionately influence the thresholds for sample characterization as resilient even when the majority of the sample population shows only normative patterns of coping under stress. Our goal in this paper is to present methods to identify these "influencers" and to control for their impact during model estimation. This technique decreases the likelihood to characterize a population as resilient when there is little evidence of exceptional performance by most individuals within the sample. The proposed influencer-adjusted residual approach to modeling resilience results in more adequate predicted outcome values and residuals than other statistical approaches. Conceptionally, however, this approach to data analysis cannot resolve the debate over whether the threshold for being characterized as resilient should be based on an entire study sample or on a subsample with influencers extracted.

Keywords Quantification · Residualization · Influence statistics · Threshold · Resilience

Introduction

Experiences that are described as adverse, traumatic, or stressful are commonly expected to negatively affect human wellbeing and development. Expected outcomes can be modeled as positive linear dose–response relationships: the worse the experience, the worse the outcome (McLaughlin et al., 2010; Sameroff, 2000). However, when different individuals are faced with the same adversity such as abuse, poverty, or acts of war, they are not all affected in the same way, i.e., there is multifinality to our responses to adverse experiences (Cicchetti & Rogosch, 1996), including many different patterns of resilience (Bonanno, 2004; Carver, 1998; Ungar, 2016). Generally, there is a large proportion of any exposed population that shows behaviors ranging from maladaptive coping to more normative patterns of recovery or adaptation to a stressor. Typically, a smaller proportion of these stress exposed samples demonstrate better than expected patterns of coping (i.e., they achieve developmental or psychological outcomes that are exceptional for the overall sample) given what they have experienced (Ungar, 2019). This results in a continuum from maladapted on one side, to normative (i.e., the expected effect of adverse experiences) in the center, and resilient on the other side (Fig. 1). This disaggregation of the sample into those that meet culturally defined norms for stress responsivity and those that surprise researchers with their ability to engage in patterns of coping that "beat the odds" weighted against them means that any single population requires a nuanced approach to understanding which individuals use which strategies to cope best. Hence, the reference point or threshold for where any one individual is on the maladapted-resilient continuum depends on the statistical contribution each individual makes when modeling a study sample (Fig. 1). Residualization, i.e., a statistical method to quantify resilience as an outcome, has been applied in scientific studies of resilience (Table 1) to derive sample-specific thresholds that have been used to situate individuals on this continuum.

Statistically, however, the right amount of a few individuals is sufficient to significantly change either the magnitude of estimated maladaptation or resilience of other individuals, or if

✉ J. Höltge
j.hoeltge@protonmail.com

[1] School of Social Work, Resilience Research Centre, Dalhousie University, 6420 Coburg Rd, Halifax, NS B3H 4R2, Canada

Published online: 29 September 2022

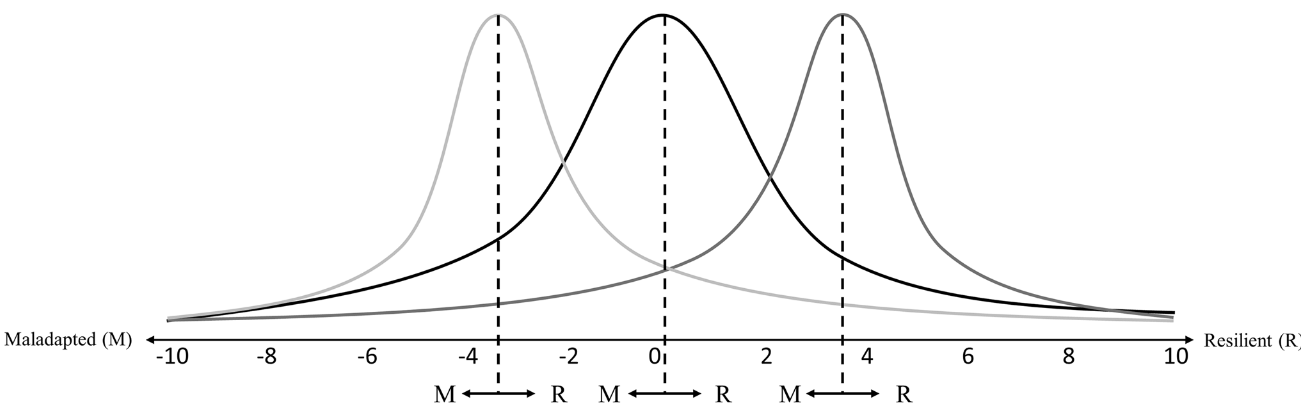


Fig. 1 Illustration of how samples can differ in their reference point that categorizes individuals as showing maladaptive, normative, or resilient patterns of coping. *Note*. The middle curve shows an equal distribution of maladaptive and resilient individuals. Sample characteristics could lead to a higher proportion of individuals showing maladaptation (right curve) or resilient behaviors (left curve)

Table 1 Exemplary studies that have used the residual approach

Author	Sample	Predictor	Outcome
Amstadter et al., 2014	$N = 7500$ adult twins	Sum of stressful life events during past year	Internalizing symptoms during past month
Bowes et al., 2010	$N = 1116$ child twins	Level of bullying victimization during primary school	Emotional and behavioral problems
Collishaw et al., 2016	$N = 262$ adolescents	Index for parent illness-related risks	Mood disorder symptom counts and behavior disorder symptom counts
van Harmelen et al., 2017	$N = 2389$ adolescents	Index of childhood family experiences	Index of psychosocial functioning
Kim-Cohen et al., 2004	$N = 1116$ child twins	Socioeconomic deprivation	Antisocial behavior and IQ
Miller-Lewis et al., 2013	$N = 474$ children	Cumulative family adversity	Mental health difficulties
Sapouna & Wolke, 2013	$N = 3136$ adolescents	Level of bullying victimization	Depression and delinquency
Veer et al., 2021	$N = 15{,}970$ adults	Stressor exposure (general stressors, Corona crisis-related stressors)	Changes in internalizing symptoms

other individuals are characterized as maladapted, normative, or resilient in the first place. This could unintentionally result in resilient individuals appearing to be more common than they really are (Masten, 2001). This potential bias in how individuals are characterized based on outcome definitions of resilience, i.e., showing better wellbeing and development in adverse contexts (Kalisch et al., 2017; Luthar & Cicchetti, 2000), has rarely been taken into account. The aim of this paper is, therefore, to describe a method for the more accurate statistical quantification of resilience as an outcome. This paper will start with an overview of the residual approach and related research before introducing influential statistics that can help to advance the residual approach for better model estimations in future studies. The statistics are then illustrated and compared using simulated data.

Residualization

Traditionally, residualization is used to take care of multicollinearity (York, 2012). Multicollinearity appears in situations where predictors of the same outcome highly correlate with each other which leads to inflated standard errors. This limits the possibility of a multiple regression to correctly estimate the independent effects of the predictors. Residualization is applied in these situations to residualize the predictors with the most inflated standard errors and to use their residualized version in the regression model which eliminates collinearity (York, 2012).

From a resilience perspective, the residual approach works as follows: an outcome such as disability, functionality, or wellbeing is regressed on an indicator of experienced adversity such as an event list or the perceived stress during a specific timespan. These studies define the resulting regression line as the expected effect of adverse experiences on an outcome (see Table 1 for an overview of studies that have applied this method). Hence, it is the reference point or line for what is normative/average and what is not: the closer/farther away an individual's observed effect is from its expected effect, the more/less the individual corresponds to what would be the expected outcomes for this individual, which is quantified via residuals. Residuals are the statistical error: they indicate the difference between what would be expected by the model and an individual's actual measurement (Field, 2018). The slope of the regression line is expected to be negative when a positive outcome such as life satisfaction is used such that individuals with a negative residual are maladapted while individuals with a positive residual (i.e., they show more life satisfaction than expected given the level of experienced adversity) are more resilient than the norm (Fig. 2A). A positive slope is expected when using a negative outcome such as depression: individuals with a negative residual (i.e., they show less depression than expected given the level of experienced adversity) are more resilient, and individuals with a positive residual are more maladapted than the norm (Fig. 2B). In a second step, these residuals are then regressed on diverse variables to investigate what predicts maladaptation, normative coping, and resilience.

The validity of this approach, however, is highly dependent on the sample and especially on those individuals within the sample who have a strong influence on the slope of the regression line (Fig. 3). To illustrate, in Fig. 3A, all individuals measure as expected, with no single individual's score influencing the statistical model more than another. Figure 3B, meanwhile, shows how influential just one individual can be when that individual's outcome is exceptionally atypical of the total sample, leading to significantly altered characterization of others in the sample. Conservatively speaking, there remains only one individual who behaves as expected, while six are now reclassified into the maladapted group and three into the resilient group depending on the exceptional functioning of the single case. This strong influence of one person can also be seen in the significant change in R^2. Figure 3C extrapolates from Fig. 3B by introducing two more influential individuals who change the R^2 to almost zero which would indicate that the model is overall statistically insignificant even though clearly many individuals within the sample show distinctly different patterns of adaptation to adversity.

Hence, we question whether the standard residual approach is adequate for determining who is maladapted, normative, or resilient if just a relatively small number of individuals can alter how the majority of a sample is being categorized. Extreme cases, as shown on Fig. 3B and 3C, are known to significantly influence residuals and lead to false predictions and therefore expectations of how the overall sample will perform (Field, 2018; Osborne & Overbay,

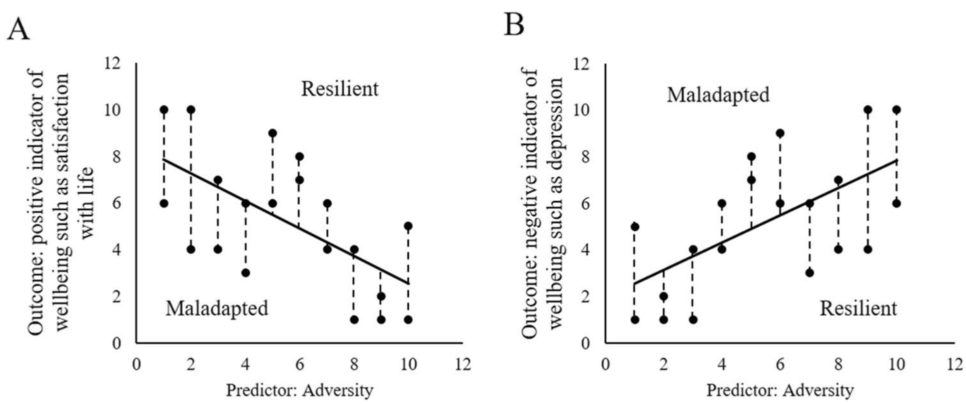


Fig. 2 Illustration of residualization for a (**A**) positive and (**B**) negative outcome. *Note*. The dashed lines represents the distance to the regression line which is commonly known as the residual

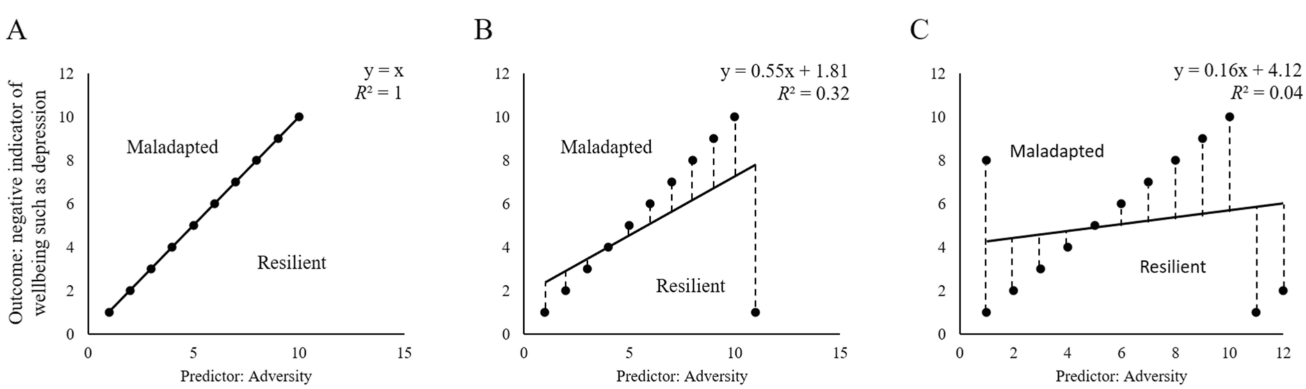


Fig. 3 The influence of extreme cases on the reference for being more or less resilient. *Note*. The dashed line represents the distance to the regression line which is commonly known as the residual

2004). Not every extreme case, though, affects the regression parameters and residuals. The ones that have a significant effect on the predictions of a model are called influencers.

The Bias of Outliers, High Leverage, and Influencers

There are three types of extreme cases: (1) an individual who does not follow the general trend in the outcome (shows an unexpected outcome value) but is in the expected range of the predictor is an outlier (Fig. 4A); (2) an individual who shows unexpected values in the predictor but follows the general trend in the outcome has high leverage (Fig. 4B); and (3) an individual who shows unexpected values in the predictor (i.e., high leverage) and does not follow the general trend in the outcome is an influencer (Fig. 4C). All three can significantly change model estimates depending on their magnitude. Outliers usually affect only the intercept of the regression, while influencers significantly alter the intercept and slope. A high leverage can exert a similarly significant impact like that of influencers since they pull in the opposite direction of influencers (Field, 2018).

Significant influencers (Fig. 4C) show unexpected, non-normative scores in both the predictor and outcome variables, such as very low stress and very low functionality (i.e., individuals who show very little resilience/pronounced maladaptation) or very high stress and very high functionality (i.e., highly resilient individuals). Hence, their influence will be the highest, on the intercept and slope of the regression line. However, outliers (Fig. 4A) who have approximately average scores on the predictor but unusual scores on the outcome will show small influence on the overall statistical model. Hence, Fig. 3 shows that the influence of individuals on a model's estimates increases the farther away they are from the mean of both the predictor and outcome.

Nevertheless, how strongly these extreme cases affect the results of a regression is also related to the size of the studied sample (Osborne & Overbay, 2004). On the one hand, it is more likely that a (normally distributed) dataset that gets closer to representing a wider population includes influencers, high leverage, and outliers, because collecting more participants makes it more likely that more of these will also be captured. On the other hand, the influence of these cases will get smaller the larger and more representative the sample (Osborne & Overbay, 2004). Hence, even though it is generally recommended to do these influencer analyses as part of the data preparation before the actual data analysis (Field, 2018; Osborne & Overbay, 2004), it can be expected that the effects of extreme cases decrease with an increasing size and representativeness of the sample.

In summary, the validity of the residual approach as it has been used so far in resilience research depends on the presence of influencers. Significant influencers are characterized by strong deviations from the normative sample's expected effect of the predictor on the outcome and significantly influence each individual's reference point for being characterized as maladapted, normative, or resilient. They shape the effect in an unexpected direction compared to the effect of the majority of a sample. If the effect of influencers is not considered, predictions can be significantly misdirected which will lead to inaccurate residuals and characterizations of being more or less resilient. As Figs. 3 and 4 show, the standard residual approach is adequate for classifications around the mean of the predictor and outcome. However, when significant influencers are present, the residuals of individuals with an increasing leverage but who are in line with the general trend of the outcome will likely be overestimated, and the residuals of significant influencers will likely be underestimated since they significantly pull the regression line towards them which minimizes their residuals (Fig. 4C).

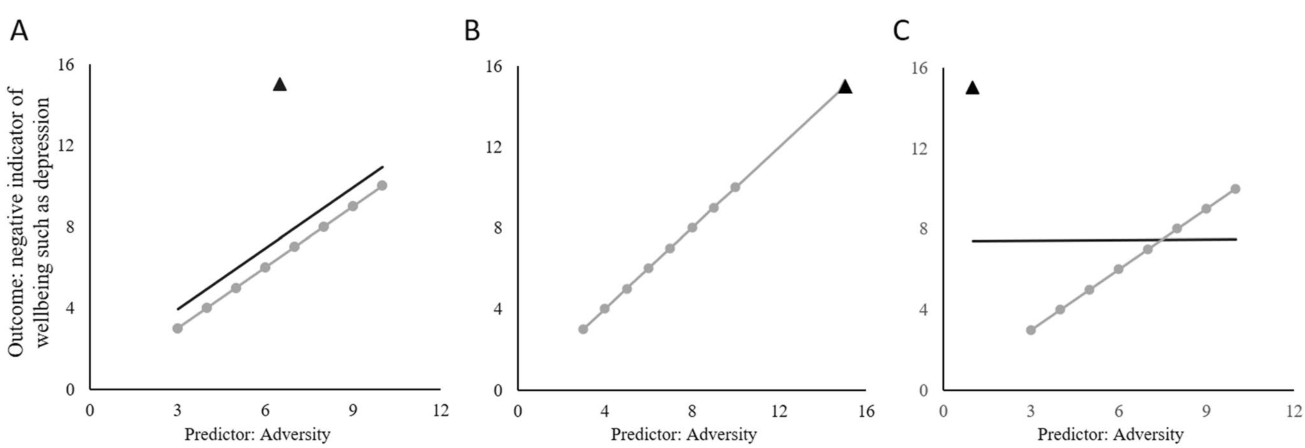


Fig. 4 Difference between an outlier (unexpected measured outcome value within expected predictor range) (**A**), high leverage (unexpected measured predictor value within trend of outcome) (**B**), and an influencer (unexpected measured predictor and outcome values) (**C**). *Note*. Grey line is the resulting model without the influence of the extreme case. Black line is with the influence of the extreme case

Research Implications

The available influence statistics (see Field, 2018) provide two options to deal with the limitations of the residual approach in the presence of significant influencers. Basically, influence statistics quantify how the prediction, expectation, or reference line differs when an individual is included vs. excluded from the sample that is used to estimate the model.

Option 1: The Individual Approach Using Deleted Studentized Residuals

Similar to using standardized residuals as has been done in past research, the deleted studentized residual (DSR) is a standardized measure that indicates the distance between an individual's observed value and their predicted value by the model (the predicted value is defined as what would be expected and normative for that individual in a specific context given their experience of adversity). However, the DSR uses the adjusted predicted value. It results from predicting the outcome value for an individual when the individual itself was excluded from the model estimation. The adjusted predicted value will be equal to the unadjusted predicted value if that individual had no influence on the regression parameters. Hence, this approach is also an indicator for the stability of the model.

Option 2: The Sample Approach Using Influencer-Adjusted Residuals

There is one obvious issue with the DSR: it only corrects for the influence of the one individual for whom the adjusted predicted value is estimated. Hence, all other individuals of a sample, including the influencers, are still present during the model estimation and significantly influence the adjusted predicted value of the respective individual. Also, the DSR will not show much difference in the case of an influencer if other influencers with similar unexpected scores are present.

An alternative approach, therefore, is needed to correct for all significant influencers at once, i.e., it needs to exclude all influencers when estimating the model for the sample, but it also needs to address the possibility of deriving predictions which are caused by the influencers. This reasoning leads to the following steps for data analysis:

1. Identify significant influencers.
2. Estimate the model with all significant influencers excluded.
3. Use the resulting model to derive the residuals for the entire sample including the influencers.

Hence, this approach advocates for keeping individuals with unexpected or unusual values in the dataset, because studying their patterns is essential for a deeper insight into resilience processes. In the next section, we demonstrate the similarities and differences between using standardized residuals, deleted studentized residuals, and influencer-adjusted residuals to quantify resilience as an outcome.

Influence Statistics

Several R packages are available to identify outliers, high leverage, and influencers such as car (Fox & Weisberg, 2019) and olsrr (Hebbali, 2020). Even though several influence statistics exist (see Field, 2018, for an overview), DFBeta and Cook's distance are the most useful when doing residualization in a resilience context.

DFBeta and Cook's distance can be used to identify cases that significantly influence the regression parameters and predicted values for all individuals of a dataset (Field, 2018). DFBeta shows this for each respective parameter, i.e., the intercept and estimate of each predictor in the equation. For each individual, DFBeta indicates how much each parameter of the regression model differs between the model with all cases included and the model that is estimated without that individual. Hence, DFBeta might rather be applied in cases where covariates are included in the model, but there is only interest in how much an individual case influences the estimate for the experienced adversity. Cook's distance can account for the overall influence of an individual on all model parameters. This would be the preferred metric when applying the residual approach with one predictor. The advantage over DFBeta in this case would be that Cook's distance incorporates the influence on the intercept and regression coefficient for the experienced adversity, while DFBeta only applies to one of them. However, influential cases as well as individuals with a high leverage in the presence of significant influencers usually affect both, and outliers usually affect the intercept. And because the residual approach has been used so far with only one predictor in the model that indicates the severity of experienced adversity, Cook's distance should therefore be the preferred influence measure when residualization is used in a resilience context.

Tutorial with Simulated Cross-Sectional Data

To illustrate the significant effect that influencers can have on model estimation, a small, simulated cross-sectional dataset was used (Table 2). The present analysis was done using the R package *olsrr v0.5.3* (Hebbali, 2020), and the respective step-by-step R syntax can be found in the supplementary material for applying it to other datasets.

Figure 5A demonstrates how the model looks when all individuals from the dataset are included in the estimation process without any adjustment: the slope of the regression line is small, and almost all cases have a residual and would be characterized as more or less resilient. The respective regression model is $y = 6.12 + 0.18x$. However, the relationship between the predictor and outcome would be exactly positive and linear if the four influential cases, especially the influencers and the case with high leverage, were not included in the model. Further, all the cases that would lie on the regression line would have no residual and would, therefore, be considered normative. Nevertheless, the farther away these normative cases are from the mean of the predictor and outcome when the influencers are present, the larger their residuals become. Thus, the residuals of the influential cases are smaller than they would be compared to the regression line that would result without their influence.

Therefore, the visual inspection of Fig. 5A shows the need to conduct an influence analysis. Because this example only used one predictor without any covariates to be in line with past empirical studies, Cook's distance was used to detect influential cases. In order to conduct this analysis, a linear or polynomial regression model (whichever fits best to the data) needs to be estimated first that only includes a predictor, which indicates a person's level of experienced risk, and an outcome, which is either a positive or negative indicator for an individual's level of wellbeing, functioning, health, etc.

As expected, Cook's distance identified the two influencers as well as the case with high leverage as significant

Table 2 Statistics for a simulated sample to demonstrate the differences between the different residual approaches

Individual	Predictor	Observed Outcome	Predicted outcome$_{ua}$	Predicted outcome$_{ia}$	Predicted outcome$_{DSR}$	Residual$_{ua}$	Residual$_{ia}$	ZResidual$_{ua}$	ZResidual$_{ia}$	DSR
1 (influencer)	1.00	16.00	6.29	1.86	4.14	9.71	14.14	1.82	2.23	2.47
2	1.00	1.00	6.29	1.86	7.47	−5.29	−.86	−.99	−.11	−1.16
3	2.00	2.00	6.47	2.86	7.23	−4.47	−.86	−.84	−.11	−.94
4	3.00	3.00	6.65	3.86	7.13	−3.65	−.86	−.69	−.11	−.74
5	4.00	4.00	6.83	4.86	7.12	−2.83	−.86	−.53	−.11	−.56
6	5.00	5.00	7.00	5.86	7.18	−2.00	−.86	−.38	−.11	−.39
7 (outlier)	5.50	15.00	7.09	6.36	6.44	7.91	8.64	1.49	1.38	1.72
8	6.00	6.00	7.18	6.86	7.27	−1.18	−.86	−.22	−.11	−.23
9	7.00	7.00	7.36	7.86	7.38	−.36	−.86	−.07	−.11	−.07
10	8.00	8.00	7.53	8.86	7.49	.47	−.86	.09	−.11	.09
11	9.00	9.00	7.71	9.86	7.58	1.29	−.86	.24	−.11	.25
12	10.00	10.00	7.89	10.86	7.62	2.11	−.86	.40	−.11	.42
13 (high leverage)	15.00	15.00	8.77	15.86	5.87	6.23	−.86	1.17	−.11	1.55
14 (influencer)	16.00	1.00	8.95	16.86	13.81	−7.95	−15.86	−1.49	−2.45	−2.25

ua unadjusted, *ia* influencer-adjusted, *DSR* deleted studentized residual, *ZResidual* standardized residual

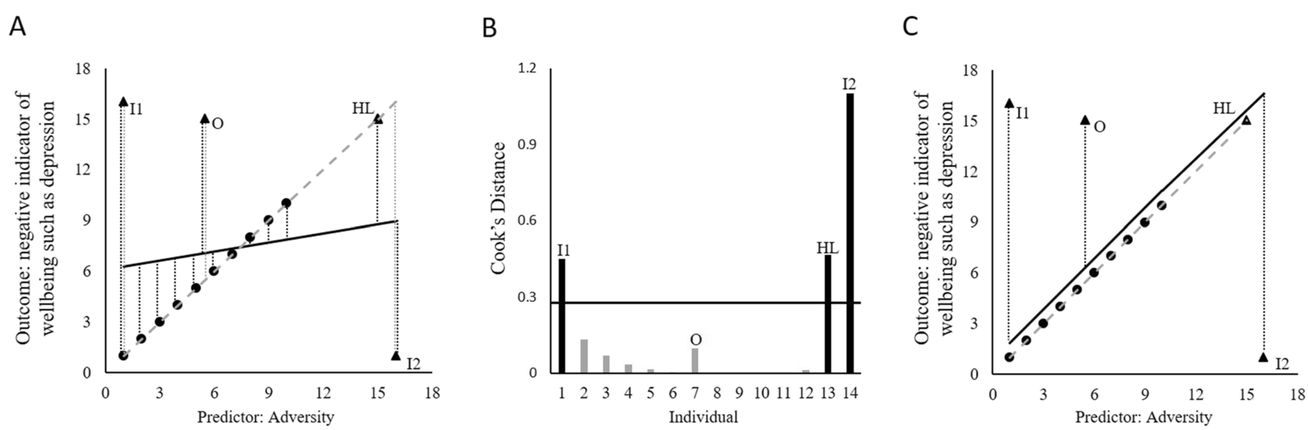


Fig. 5 Effect of the influencer-adjusted approach. *Note*. (**A**) The black, solid slope line represents the resulting model if all data would be used. The grey, dashed slope line represents the model if none of the potential influential cases would be used. The dotted, black lines represent the residuals for the model based on the entire dataset. The dotted, grey lines represent the residuals for the model based on the dataset without the potential influential cases. (**B**) Graphical representation of Cook's distance for each individual. Cases above the threshold (.287) can be considered influential cases. (**C**) The black slope line represents the resulting model when the identified influencers based on Cook's distance (i.e., I1, I2, HL) are excluded. The grey, dashed slope line represents the model without the influence of any influential case. I(1,2): influencer, O: outlier, and HL: high leverage

influencers of the model (Fig. 5B). Hence, these three cases were excluded in the third step to estimate the influencer-adjusted model via linear regression again, which resulted in $y = 0.86 + 1x$. In the fourth and last step, this model was used to predict each individual's outcome value and to derive the individual residuals. As Fig. 5C shows, the resulting model is very close to the normative model, with only the outlier increasing the intercept by 0.86.

Table 2 shows the accuracy of the predicted values for the three approaches, i.e., using unadjusted residuals, deleted studentized residuals, or influencer-adjusted residuals. Overall, the influencer-adjusted approach predicts the outcome values for all individuals, including influencers and the individual with high leverage, in relation to their given predictor value more accurately than the other two approaches when considering the trend of the majority of the sample. For example, following the trend of the sample, a person with a predictor/risk score of 3 should likely have an outcome score around 3. While using unadjusted residuals predicts an outcome score of 6.47 and using deleted studentized residuals predicts a score of 7.23, using influencer-adjusted residuals predict a score of 2.86. Further, as the unstandardized residuals show, the unadjusted approach underestimates the residuals for the influencers and overestimates the residuals for most of the normative cases. In contrast, the influencer-adjusted residuals represent the residuals more precisely when the reference point for being maladapted, normative, or resilient is based on the normative sample: the influencers and outlier have large residuals compared to the rest of the sample that is very close to the regression line.

Longitudinal Studies: Thresholds of Resilience over Time

The influencer-adjusted residual approach has been presented based on cross-sectional data. The discussion about the correct method to differentiate between maladapted, normative, and resilient individuals becomes more complex in longitudinal studies where adversity is an already present contextual variable at the beginning of a study such as poverty or chronic illness. The options are as follows:

1. Use the baseline threshold for all subsequent timepoints.
2. Use timepoint-specific thresholds.
3. Use the average threshold across all timepoints as the threshold for all timepoints.

One of the advantages of the first option is the use of the baseline sample and, therefore, the likelihood that the dataset with the richest information and highest variability will be used despite an anticipated decline in participant numbers during longitudinal studies. By using the largest sample possible, the resulting threshold for being characterized as maladapted, normative, or resilient would be closest to a population parameter for that sample. In reference to studies that make aware of the higher likelihood of attrition for vulnerable participants (e.g., Kim et al., 2014; Rotherbühler & Voorpostel, 2016), a significant dropout could likely occur in resilience studies for maladapted individuals. Such a change of the sample over time would lead to biased timepoint-specific thresholds at subsequent timepoints. This dynamic can make individuals who have been characterized as resilient at baseline into maladapted individuals at subsequent timepoints just because the threshold for being characterized as maladapted has been lowered due to the dropout of highly maladapted individuals. Without such a dropout, these individuals would still be considered resilient at subsequent timepoints.

Consequently, the second option, i.e., using timepoint-specific thresholds, would be most appropriate in longitudinal studies without dropout and should be accompanied by an analysis of the variability of the samples' threshold over time. A significant deviation at a certain timepoint could hint to more macro-scale events that influence the resilience of the majority of a sample. If the second option is used in longitudinal studies with dropout, influence statistics need to be applied at each timepoint to prevent biased thresholds due to the dropout of certain influential subpopulations.

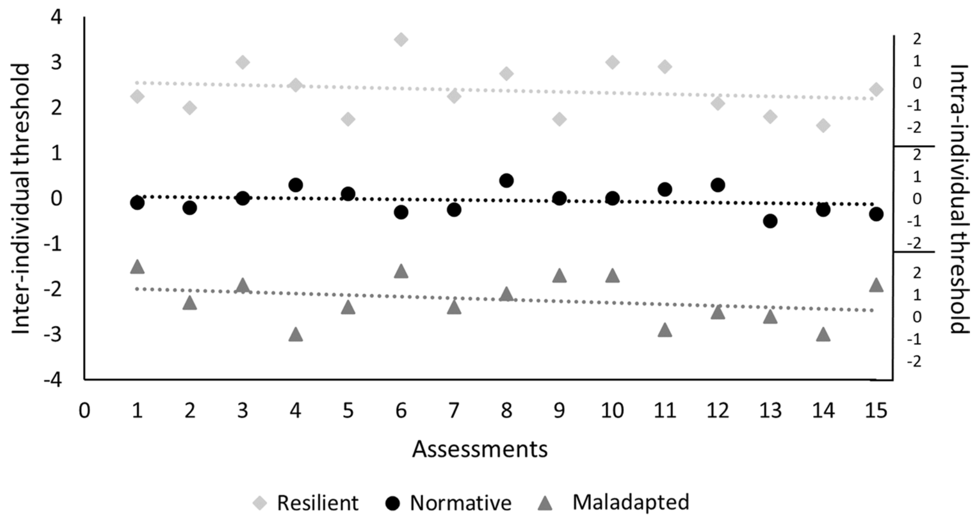
The advantage of the third option would be a threshold that is generalizable over time. However, the same decision needs to be made in relation to cross-sectional data: if thresholds are present at certain timepoints that significantly differ from the majority of the thresholds associated with the other timepoints, should these exceptional, influential thresholds be excluded from calculating the average threshold across time?

Finally, time-intensive longitudinal studies give the possibility to derive intra-individual thresholds. Hence, an individual can be categorized as maladapted, normative, or resilient by using one of the above-described thresholds that are based on the entire sample, and the individual can be characterized by the deviations from their own threshold over time (Fig. 6). This can give an insight into what characterizes an individual when they behave expected vs. unexpected in relation to their own threshold which can be compared between different individuals with similar and contrasting patterns (e.g., comparing liable with stable individuals). These intra-individual dynamics might also be important to explain why some individuals are able to transition between being maladapted, normative, and resilient and others who show a stable trajectory of maladaptation or resilience.

Reporting the Results of an Influencer Analysis

Future studies should provide the name of the influence statistic that has been used and why it has been chosen and show the regression model with all cases included and the regression model with significant influencers, high leverage, and

Fig. 6 Inter-individual and intra-individual thresholds in longitudinal studies. *Note*. The characterization as resilient, normative, and maladapted is based on the inter-individual threshold



outliers removed (see Table 3 for an overview of the central constructs). In the case of longitudinal studies, researchers need to justify why they chose which kind of threshold.

Theoretical Considerations: What is Normative and Unexpected?

Statistically, the impact of influencers needs to be controlled for when studying resilience processes using residualization if more accurate residuals are to be achieved for the majority of individuals in a sample. As presented, one way to do this is to exclude influencers from determining thresholds (i.e., the model estimation) that decide who shows high and low resilience. Conceptually, however, the main question that arises when using influence statistics is to decide what is normative? Should the threshold or reference point for characterizing someone's behavior as maladapted, normative, or resilient be based on the entire sample including influencers, or should this determination be based on the sample without its influencers? Setting the threshold without influencers would make the influencers more or less resilient (i.e., more

Table 3 Overview of central constructs

Construct	Definition
Resilience	Showing better-than-expected wellbeing and development in adverse contexts
Residualization	The regression line indicates the expected level of an outcome in relation to different levels of adversity. The closer/farther away an individual's observed effect is from its expected effect, the more/less the individual corresponds to what would be the expected outcomes for this individual, which is quantified via residuals
Extreme cases	
Outlier	An individual who does not follow the general trend in the outcome (shows an unexpected outcome value) but is in the expected range of the predictor
High leverage	An individual who shows unexpected values in the predictor but follows the general trend in the outcome
Influencer	An individual who shows unexpected values in the predictor (i.e., high leverage) and does not follow the general trend in the outcome (i.e., outlier)
Residuals	
Unadjusted	Difference between the predicted and observed outcome scores for an individual
Deleted studentized	Difference between the adjusted predicted and observed outcome scores for an individual. The adjusted predicted outcome score results from excluding only the individual from the model estimation for whom the residual is being calculated
Influencer-adjusted	Difference between the influencer-adjusted predicted and observed outcome scores for an individual. The model is estimated without any influencers included and then used to estimate the residuals for all individuals
Influence measures	
DFBeta	Results in estimates for each respective parameter of the regression model. Useful when covariates are included in the model but only one variable is of interest
Cook's distance	Accounts for the overall influence of an individual on all model parameters. Useful when influencers need to be identified and only one variable is present in the model

unexpected) and everyone else more normative/expected. Using the entire sample, meanwhile, leads to the opposite effect and makes a greater proportion of the sample as a whole appear to score higher or lower on resilience than they might merit on their own.

This nuanced examination of resilience data facilitates a more accurate portrayal of a sample and ensures that those who are coping in normative ways are not wrongly classified as exceptional and that those who are resilient are not overlooked. The more accurate we are with our predictions, the better we can inform policy and practice by (1) focusing on the ways truly exceptional individuals manage to cope with adversity, learning from them, and documenting their strategies and (2) not overusing the term resilience to label normal developmental processes within a population as exceptional when they are in fact within the common range of human experience in that specific context. A better understanding of influencers in statistical models can help to achieve both these goals.

Words of Caution

As Fig. 3C indicates, a specific sample composition can lead to a non-significant effect of the predictor on the outcome by resulting in a horizontal regression line with an explained variance of zero. Nevertheless, the residual approach will categorize each individual as maladapted, normative, or resilient due to the difference between the statistically expected and empirically measured outcome values. However, is it realistic to expect all individuals to show about the same outcome value at any given stressor severity and use this value as the reference point? Published resilience studies that have applied residualization show explained variances of 1.7–7% (Miller-Lewis et al., 2013), 3% (Amstadter et al., 2014), 4–21% (Veer et al., 2021), and 24–29% (van Harmelen et al., 2017). As this paper has shown, low explained variances can be due to influencers and therefore point to the need to control for them. However, influencers, outliers, and cases with high leverage altogether can also be responsible for an estimated significant effect, and excluding them from the model estimation can lead to a non-significant effect and lower explained variance. This leads to the following conceptual questions: is there no estimated effect because of the choice of predictor and outcome and, therefore, the underlying theoretical model might not be correct, or is a very low explained variance and effect a sign of a pronounced multifinality at any level of adversity? Another possibility could be a bimodal distribution of a sample, i.e., almost distinct subgroups of low and highly resilient individuals who have almost exactly opposite values and should, therefore, receive their own respective thresholds of resilience (Field, 2018). Such a sample composition often results in a low explained variance. In such cases, person-centered approaches such as latent profile analysis should be applied to identify these subgroups before residualization is used.

Another open conceptual question in the application of the residual approach in resilience research concerns the missing of a “region of normativity or expectation”: where does normativity end and where does the unexpected start? Statistically speaking, at which point is the individual effect of the adversity on the outcome significantly different from the normative expected effect? So far, normativity has been quantified as the exact match between an individual’s observed outcome score and its predicted score by the statistical model, and any deviation in either direction (a positive or negative residual) has been used as an indicator for how unexpected an individual is. However, this is similar to the past application of mixed effect models where it was mainly of interest to investigate if there is a random effect at all without going deeper and looking for who is significantly different from the average and who is not (Rodriguez, Williams, & Rast, 2021). Rodriguez et al. (2021) introduced a statistical method to identify individuals who significantly differ from the average effect of the sample. Therefore, this method could be used for the identification of a region of normativity in resilience research after an adequate model has been estimated based on influence statistics. Future studies are asked to investigate if normativity should depend on a point estimate or if there rather exist a tolerance for being normative which would in turn affect the proportion for being characterized as maladapted and resilient.

Conclusion

Resilience researchers need to be aware that how un-/expected an individual’s resilience is evaluated in a stressful context is determined by a threshold that is sensitive to the sample on which it is based. Statistically, this evaluation can be very misguided in a sample that includes influencers who significantly determine this threshold in cross-sectional and longitudinal studies. Hence, researchers need to first identify significant influencers and control for their impact before they quantify the resilience (as an outcome) of an individual. The proposed influencer-adjusted residual approach serves to relativize how a normative response in a stressful context should look. It situates individuals who have reported normative values closer to the expected norm and exceptional cases farther from the norm. Nevertheless, while this paper recommends using the influencer-adjusted approach, it remains a conceptual debate about who should decide about a sample’s normative resilience threshold: everyone including influencers or the norm without influencers?

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s42844-022-00078-6.

Declarations

Conflict of Interest The authors declare no competing interests.

References

Amstadter, A. B., Myers, J. M., & Kendler, K. S. (2014). Psychiatric resilience: Longitudinal twin study. *The British Journal of Psychiatry, 205*(4), 275–280. https://doi.org/10.1192/bjp.bp.113.130906

Bonanno, G. A. (2004). Loss, Trauma, and human resilience: Have we underestimated the human capacity to thrive after extremely aversive events? *American Psychologist, 59*(1), 20–28. https://doi.org/10.1037/0003-066X.59.1.20

Bowes, L., Maughan, B., Caspi, A., Moffitt, T. E., & Arseneault, L. (2010). Families promote emotional and behavioural resilience to bullying: Evidence of an environmental effect. *Journal of Child Psychology and Psychiatry, 51*(7), 809–817. https://doi.org/10.1111/j.1469-7610.2010.02216.x

Carver, C. S. (1998). Resilience and thriving: Issues, models, and linkages. *Journal of Social Issues, 54*(2), 245–266. https://doi.org/10.1111/j.1540-4560.1998.tb01217.x

Cicchetti, D., & Rogosch, F. A. (1996). Equifinality and multifinality in developmental psychopathology. *Development and Psychopathology, 8*(4), 597–600. https://doi.org/10.1017/S0954579400007318

Collishaw, S., Hammerton, G., Mahedy, L., Sellers, R., Owen, M. J., Craddock, N., Thapar, A. K., Harold, G. T., Rice, F., & Thapar, A. (2016). Mental health resilience in the adolescent offspring of parents with depression: A prospective longitudinal study. *The Lancet Psychiatry, 3*(1), 49–57. https://doi.org/10.1016/S2215-0366(15)00358-2

Field, A. P. (2018). *Discovering Statistics Using IBM SPSS Statistics* (5th ed.). Sage.

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression* (3rd ed.). Sage.

Hebbali, A. (2020). *olsrr: Tools for Building OLS Regression Models. R package version 0.5.3*. https://CRAN.R-project.org/package=olsrr

Kalisch, R., Baker, D. G., Basten, U., Boks, M. P., Bonanno, G. A., Brummelman, E., Chmitorz, A., Fernàndez, G., Fiebach, C. J., Galatzer-Levy, I., Geuze, E., GroppaHelmreich, S. I., Hendler, T., Hermans, E. J., Jovanovic, T., Kubiak, T., Lieb, K., Lutz, B., & Kleim, B. (2017). The resilience framework as a strategy to combat stress-related disorders. *Nature Human Behaviour, 1*(11), 784–790. https://doi.org/10.1038/s41562-017-0200-8

Kim, R., Hickman, N., Gali, K., Orozco, N., & Prochaska, J. J. (2014). Maximizing retention with high risk participants in a clinical trial. *American Journal of Health Promotion, 28*(4), 268–274. https://doi.org/10.4278/ajhp.120720-QUAN-355

Kim-Cohen, J., Moffitt, T. E., Caspi, A., & Taylor, A. (2004). Genetic and environmental processes in young children's resilience and vulnerability to socioeconomic deprivation. *Child Development, 75*(3), 651–668. https://doi.org/10.1111/j.1467-8624.2004.00699.x

Luthar, S. S., & Cicchetti, D. (2000). The construct of resilience: Implications for interventions and social policies. *Development and Psychopathology, 12*(4), 857–885. https://doi.org/10.1017/S0954579400004156

Masten, A. S. (2001). Ordinary magic: Resilience processes in development. *American Psychologist, 56*(3), 227–238. https://doi.org/10.1037//0003-066X.56.3.227

McLaughlin, K. A., Conron, K. J., Koenen, K. C., & Gilman, S. E. (2010). Childhood adversity, adult stressful life events, and risk of past-year psychiatric disorder: A test of the stress sensitization hypothesis in a population-based sample of adults. *Psychological Medicine, 40*(10), 1647–1658. https://doi.org/10.1017/S0033291709992121

Miller-Lewis, L. R., Searle, A. K., Sawyer, M. G., Baghurst, P. A., & Hedley, D. (2013). Resource factors for mental health resilience in early childhood: An analysis with multiple methodologies. *Child and Adolescent Psychiatry and Mental Health, 7*(1), 1–23. https://doi.org/10.1186/1753-2000-7-6

Osborne, J. W., & Overbay, A. (2004). The power of outliers (and why researchers should always check for them). *Practical Assessment, Research, and Evaluation, 9*(1), 6. https://doi.org/10.7275/qf69-7k43

Rodriguez, J. E., Williams, D. R., & Rast, P. (2021). *Who Is and Is Not "Average"? Random Effects Selection with Spike-and-Slab Priors*. PsyArXiv. https://doi.org/10.31234/osf.io/4d9tv

Rothenbühler, M., & Voorpostel, M. (2016). Attrition in the Swiss Household Panel: Are vulnerable groups more affected than others?. In M. Oris, C. Roberts, D. Joye, & M. E. Stähli (Eds). *Surveying Human Vulnerabilities across the Life Course* (pp. 223–244). Springer Nature. https://doi.org/10.1007/978-3-319-24157-9

Sameroff, A. J. (2000). Dialectical processes in developmental psychopathology. In A. Sameroff, M. Lewis, & S. Miller (Eds.), *Handbook of Developmental Psychopathology* (pp. 23–40). Kluwer Academic/Plenum Publishers.

Sapouna, M., & Wolke, D. (2013). Resilience to bullying victimization: The role of individual, family and peer characteristics. *Child Abuse & Neglect, 37*(11), 997–1006. https://doi.org/10.1016/j.chiabu.2013.05.009

Ungar, M. (2016). Varied patterns of family resilience in challenging contexts. *Journal of Marital and Family Therapy, 42*(1), 19–31. https://doi.org/10.1111/jmft.12124

Ungar, M. (2019). Designing resilience research: Using multiple methods to investigate risk exposure, promotive and protective processes, and contextually relevant outcomes for children and youth. *Child Abuse & Neglect, 96*, 104098. https://doi.org/10.1016/j.chiabu.2019.104098

van Harmelen, A. L., Kievit, R. A., Ioannidis, K., Neufeld, S., Jones, P. B., Bullmore, E., Dolan, R., Fonagy, P., Goodyer, I., NSPN Consortium. (2017). Adolescent friendships predict later resilient functioning across psychosocial domains in a healthy community cohort. *Psychological Medicine, 47*(13), 2312–2322. https://doi.org/10.1017/S0033291717000836

Veer, I. M., Riepenhausen, A., Zerban, M., Wackerhagen, C., Puhlmann, L. M., Engen, H., Köber, G., Bögemann, S. A., Weermeijer, J., Uscilko, A., Mor, N., Marciniak, M. A., Askelund, A. D., Al-Kamel, A., Ayash, S., Barsuola, G., Bartkute-Norkuniene, V., Battaglia, S., Bobko, Y., ... & Kalisch, R. (2021). Psycho-social factors associated with mental resilience in the Corona lockdown. *Translational Psychiatry*, *11*, 67. https://doi.org/10.1038/s41398-020-01150-4

York, R. (2012). Residualization is not the answer: Rethinking how to address multicollinearity. *Social Science Research, 41*(6), 1379–1386. https://doi.org/10.1016/j.ssresearch.2012.05.014